

OY	181	GGCCCTCTCCACATGCGTCAACACGTGCGCGTCTTTTGTGTGACGTTAAACAAAGGCGT	240
Db	181	GGTCTCTTCCACTGCTCTACACTGCTCCGTGTTTTGTGGACTATATACAAAGGCGT	240
OY	241	GGCGATGAGATTTTTCTTTTCAAATCCGCGAATAATCAAACTTATATATTTCTTAAGAA	300
Db	241	GGCGATGAGATTTTTCTTTTCAAATCCGCGAATAATCAAACTTATATATTTCTTAAGAA	300
OY	301	CACAGTCATTTCTGATCTTCTCATGATCTGACTTTTCCATTCOAAAATTCITAGTAGTC	360
Db	301	CACAGTCATTTCTGATCTTCTCATGATCTGACTTTTCCATTCOAAAATTCITAGTAGTC	360
OY	361	CAAACTGGGAACAGACCAACGAGAACTTTTGGTGTGCAAGTTTACCTCCGTCATTTTTTA	420
Db	361	CAAACTGGGAACAGACCAACGAGAACTTTTGGTGTGCAAGTTTACCTCCGTCATTTTTTA	420
OY	421	TTTTCAATGTATATACAGTATTTTTCATTTCCGCGACATGATACGATACGCTACCAAA	480
Db	421	TTTTCAATGTATATACAGTATTTTTCATTTCCGCGACATGATACGATACGCTACCAAA	480
OY	481	GACCAACAGGCGCATTTTAAACATCCAAACCCCAAAAATCTCTTGGGGGCTAAGATTCTCTC	540
Db	481	GACCAACAGGCGCATTTTAAACATCCAAACCCCAAAAATCTCTTGGGGGCTAAGATTCTCTC	540
OY	541	TGTTGTCACTCGGGCATTCGATCGTCTTACTCTCTTGGCTTAACATGATTTCTACCAACAG	600
Db	541	TGTTGTCACTCGGGCATTCGATCGTCTTACTCTCTTGGCTTAACATGATTTCTACCAACAG	600
OY	601	GCAGCCGAGACACACAGATGTGGAAGAAATGCTTTCTCTTAAATCAGATGTGGGTCTAGT	660
Db	601	GCAGCCGAGACACACAGATGTGGAAGAAATGCTTTCTCTTAAATCAGATGTGGGTCTAGT	660
OY	661	CTGGCATGAAATAGTAAATTTACATCTGTCAAGTCATTTTCTGGATTATTTCTTAATGTT	720
Db	661	CTGGCATGAAATAGTAAATTTACATCTGTCAAGTCATTTTCTGGATTATTTCTTAATGTT	720
OY	721	TATTTGTATGTATACACCTTACCAAAAGAACTGTACGCGGTATACGTAAAGAACAGAGGG	780
Db	721	TATTTGTATGTATACACCTTACCAAAAGAACTGTACGCGGTATACGTAAAGAACAGAGGG	780
OY	781	TGTAGTAAAGTCCCGCAGGAAAAAGGTGAACGTCAAAAGTTTTCATTTATCATCTGTATTT	840
Db	781	TGTAGTAAAGTCCCGCAGGAAAAAGGTGAACGTCAAAAGTTTTCATTTATCATCTGTATTT	840
OY	841	CTTTATTTGTTTGTCTCTTCCATTTTGGCCGAATTCCTTACACCCGTAGCCAAACCG	900
Db	841	CTTTATTTGTTTGTCTCTTCCATTTTGGCCGAATTCCTTACACCCGTAGCCAAACCG	900
OY	901	GGATGTCTTTGACAGCACTGCTGAAAAATACGTGTCATATGTAAGAAAGAGACACTGTG	960
Db	901	GGATGTCTTTGACAGCACTGCTGAAAAATACGTGTCATATGTAAGAAAGAGACACTGTG	960
OY	961	GTTAACTTCCTTAAATGCATGCGTGTGATCGGTGCATCTATTTTTTCTTTGGCAAGTCTT	1020
Db	961	GTTAACTTCCTTAAATGCATGCGTGTGATCGGTGCATCTATTTTTTCTTTGGCAAGTCTT	1020
OY	1021	CAGAAATTCCTTGTATAGTATGCTGAAGTCCCAATTTTGCACAATCTCTGTCCAGAGA	1080
Db	1021	CAGAAATTCCTTGTATAGTATGCTGAAGTCCCAATTTTGCACAATCTCTGTCCAGAGA	1080
OY	1081	CAATAGGAAAAAACAAGAGATGCTGTGACCCCAATGAAAGAGACCCCAATGTAAACAA	1140
Db	1081	CAATAGGAAAAAACAAGAGATGCTGTGACCCCAATGAAAGAGACCCCAATGTAAACAA	1140
OY	1141	TTAACTAAGGAATATTTTCAATCTCTTTGTGTGTCAGAACTCGTTAAA	1187
Db	1141	TTAACTAAGGAATATTTTCAATCTCTTTGTGTGTCAGAACTCGTTAAA	1187

[illegible]

QY 748 AGAAGTGTACCGGTGATACGTAAGAGAGAGGGGTGTAGGTAAAGTCCCGAGAAAAAGGT 807
DB 867 AGAAGTGTACCGGTGATACGTAAGAGAGAGGGGTGTAGGTAAAGTCCCGAGAAAAAGGT 926
QY 808 GAAGGTCAAGATTTTCAATATCATTTGCTGTATCTTTATTTTGTGTTCTTCCATTT 867
DB 927 GAAGGTCAAGATTTTCAATATCATTTGCTGTATCTTTATTTTGTGTTCTTCCATTT 986
QY 868 TGCCCGAATTCCTTACACCCCTGAGCCAAACCCGGATGTCCTTGAATGCAATGCTGAAA 927
DB 987 TGCCCGAATTCCTTACACCCCTGAGCCAAACCCGGATGTCCTTGAATGCAATGCTGAAA 1046
QY 928 TACTCTGTCTATGTGAAGAGAGACTCTGTGTTAACTCTTAAATGATGCTGGA 987
DB 1047 TACTCTGTCTATGTGAAGAGAGACTCTGTGTTAACTCTTAAATGATGCTGGA 1106
QY 988 TCCGTTCAATCTTTTCTTCCCTTGAAGTCTTGAATGCTTGAATGCTGGA 1047
DB 1107 TCCGTTCAATCTTTTCTTCCCTTGAAGTCTTGAATGCTTGAATGCTGGA 1166
QY 1048 GTGCCCAATTCCTCAACATCTCTGTCCAGAGCAATAGAAAAAGAACAGATGCTG 1107
DB 1167 GTGCCCAATTCCTCAACATCTCTGTCCAGAGCAATAGAAAAAGAACAGATGCTG 1226
QY 1108 TGACCCCAATTAAGAGACTCCATGTAAACAATTAAGTAAGTAATTTCAATCTCT 1167
DB 1227 TGACCCCAATTAAGAGACTCCATGTAAACAATTAAGTAAGTAATTTCAATCTCT 1286
QY 1168 TGTGTGAGAACTGCTTAAA 1187
DB 1287 TGTGTGAGAACTGCTTAAA 1306

RESULT 3
US-09-964-008-2
; Sequence 2, Application US/09964008
; Patent No. US20020156246A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: 15625 Receptor, A No. US20020156246A1el G-Protein Coupled Receptc
; FILE REFERENCE: 5800-13, 03800-111548
; CURRENT APPLICATION NUMBER: US/09/964,008
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/187,134
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-008-2

Query Match 97.58; Score 1156.8; DB 9; Length 2286;
Best Local Similarity 99.88; Pred. No. 1.2e-289;
Matches 1158; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 28 CTGAGGCTGCAATTAATCTACTACTGATGATATCAAAACCTCCAGATCAACAGTT 87
DB 182 CTGAGGCTGCAATTAATCTACTACTGATGATATCAAAACCTCCAGATCAACAGTT 241
QY 88 ATCAGGTACCAACAGAAATGCAAGCCGTGCAATCTCACTCTGCGCTGGGAAC 147
DB 242 ATCAGGTACCAACAGAAATGCAAGCCGTGCAATCTCACTCTGCGCTGGGAAC 301
QY 148 CAGTCTGTGACAGAGACTACAAATATCACAGAGTCTCTCCACTGCTACACTGT 207
DB 302 CAGTCTGTGACAGAGACTACAAATATCACAGAGTCTCTCCACTGCTACACTGT 361
QY 208 CCGTGTGTGTGTGAGACTATACAAATGCGCTGGCGATGAGATTTCTTCAATCG 267
DB 362 CCGTGTGTGTGTGAGACTATACAAATGCGCTGGCGATGAGATTTCTTCAATCG 421

QY 268 GAGTAATCAAACTTATATTTTCTTAAGACAGATCAATTTCTGATCTTCATGAT 327
DB 422 GAGTAATCAAACTTATATTTTCTTAAGACAGATCAATTTCTGATCTTCATGAT 481
QY 328 TCTGACTTTTCAATCAAAATTTCTTAGTATGCCAAACTGGGAACAGACACTGAGAC 387
DB 482 TCTGACTTTTCAATCAAAATTTCTTAGTATGCCAAACTGGGAACAGACACTGAGAC 541
QY 388 TTTTGTGTGACAGTTCACCTCCGTCATATTTTATTTACATGATATGATATTTAT 447
DB 542 TTTTGTGTGACAGTTCACCTCCGTCATATTTTATTTTACATGATATGATATTTAT 601
QY 448 CCTGGAGATGATTAATATGATGCTACAGACAGACAGCCAGGCAATTTAAACATCAA 507
DB 602 CCTGGAGATGATTAATATGATGCTACAGACAGACAGCCAGGCAATTTAAACATCAA 661
QY 508 CCCCCAAATCTCTTGGGGGCTAAGATTTCTGTGTGTATCTGCGCATTCATGTTCTT 567
DB 662 CCCCCAAATCTCTTGGGGGCTAAGATTTCTGTGTGTATCTGCGCATTCATGTTCTT 721
QY 568 ACTGCTTTGGCTTAACATGATTTGACCAACAGGAGCCGAGACAAAGATGGAAGA 627
DB 722 ACTGCTTTGGCTTAACATGATTTGACCAACAGGAGCCGAGACAAAGATGGAAGA 781
QY 628 ATGCTCTTCTTAAATCAGATTCGCTAGTCTGCGATGAATAGTAATTTACATCTG 687
DB 782 ATGCTCTTCTTAAATCAGATTCGCTAGTCTGCGATGAATAGTAATTTACATCTG 841
QY 688 TCAAGTCAATTTCTGATTAATTTCTTAATTTGATTTGATTTATACATCAATACAA 747
DB 842 TCAAGTCAATTTCTGATTAATTTCTTAATTTGATTTGATTTATACATCAATACAA 901
QY 748 AGAAGTGTACCGGTGATACGTAAGAGAGAGGGGTGTAGGTAAAGTCCCGAGAAAAAGGT 807
DB 902 AGAAGTGTACCGGTGATACGTAAGAGAGAGGGGTGTAGGTAAAGTCCCGAGAAAAAGGT 961
QY 808 GAAGGTCAAGATTTTCAATATCATTTGCTGTATCTTTATTTTGTGTTCTTCCATTT 867
DB 962 GAAGGTCAAGATTTTCAATATCATTTGCTGTATCTTTATTTTGTGTTCTTCCATTT 1021
QY 868 TGCCCGAATTCCTTACACCCCTGAGCCAAACCCGGATGTCCTTGAATGCAATGCTGAAA 927
DB 1022 TGCCCGAATTCCTTACACCCCTGAGCCAAACCCGGATGTCCTTGAATGCAATGCTGAAA 1081
QY 928 TACTCTGTCTATGTGAAGAGAGACTCTGTGTTAACTCTTAAATGATGCTGGA 987
DB 1082 TACTCTGTCTATGTGAAGAGAGACTCTGTGTTAACTCTTAAATGATGCTGGA 1141
QY 988 TCCGTTCAATCTTTTCTTCCCTTGAAGTCTTGAATGCTTGAATGCTGGA 1047
DB 1142 TCCGTTCAATCTTTTCTTCCCTTGAAGTCTTGAATGCTTGAATGCTGGA 1201
QY 1048 GTGCCCAATTCCTCAACATCTCTGTCCAGAGCAATAGAAAAAGAACAGATGCTG 1107
DB 1202 GTGCCCAATTCCTCAACATCTCTGTCCAGAGCAATAGAAAAAGAACAGATGCTG 1261
QY 1108 TGACCCCAATTAAGAGACTCCATGTAAACAATTAAGTAAGTAATTTCAATCTCT 1167
DB 1262 TGACCCCAATTAAGAGACTCCATGTAAACAATTAAGTAAGTAATTTCAATCTCT 1321
QY 1168 TGTGTGAGAACTGCTTAAA 1187
DB 1322 TGTGTGAGAACTGCTTAAA 1341

RESULT 4
US-09-964-008-4
; Sequence 4, Application US/09964008
; Patent No. US20020156246A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Gu, Wei

|||||
Db 276 GACATTTTCATCAAAATTTCTAGTATGCCAACTGGGAACAGACACAGCACTGTTT 335
QY 391 TGTGTGCAAGTACCTCCGATATTTTATTTTCAATGATATGATTTATTCCT 450
Db 336 TGTGTGCAAGTACCTCCGATATTTTATTTTCAATGATATGATTTATTCCT 395
QY 451 GGGACGATATCAATGATGCTGCTACAGAAAGCAGCAGGCCATTTTAAACATCCACC 510
Db 396 GGGACGATATCAATGATGCTGCTACAGAAAGCAGCAGGCCATTTTAAACATCCACC 455
QY 511 CAAAAATCTCTGGGGCTAAAGATTTCTCTGTTGTCATCTGGGCATTCATGTTCTACT 570
Db 456 CAAAAATCTCTGGGGCTAAAGATTTCTCTGTTGTCATCTGGGCATTCATGTTCTACT 515
QY 571 CTCTTTCCTTAATCAATGATTTGACCAAGCAGCAGGCCAGCAAGATGGAAGAAATG 630
Db 516 CTCTTTCCTTAATCAATGATTTGACCAAGCAGCAGGCCAGCAAGATGGAAGAAATG 575
QY 631 CTCTTTCCTTAATCAATGATTTGACCAAGCAGCAGGCCAGCAAGATGGAAGAAATG 690
Db 576 CTCTTTCCTTAATCAATGATTTGACCAAGCAGCAGGCCAGCAAGATGGAAGAAATG 635
QY 691 AGTCAATTTTCGATTAATTTCTTAATTTGTTATTTGATTTACCTCATTAACAAAAGA 750
Db 636 AGTCAATTTTCGATTAATTTCTTAATTTGTTATTTGATTTACCTCATTAACAAAAGA 695
QY 751 ACTGTACCGGTCATACGATGAGAGAGAGGGGTGTAGTAAAGTCCCGAGAAAAGTGAA 810
Db 696 ACTGTACCGGTCATACGATGAGAGAGAGGGGTGTAGTAAAGTCCCGAGAAAAGTGAA 755
QY 811 CGTCAAGATTTTCATTAATTCATTTGCTGATTTCTTAATTTGTTTCTCTTTCATTTTGC 870
Db 756 CGTCAAGATTTTCATTAATTCATTTGCTGATTTCTTAATTTGTTTCTCTTTCATTTTGC 815
QY 871 CCGAATTTCTTACACCTGAGACCAACCCGGGATGTTCTTTACCTGACCTGCTGAAATAC 930
Db 816 CCGAATTTCTTACACCTGAGACCAACCCGGGATGTTCTTTACCTGACCTGCTGAAATAC 875
QY 931 TCTGTCTATGTGAAGAAGAGACACTGTGTAACTTCTTAATGATGCTGAGATTC 990
Db 876 TCTGTCTATGTGAAGAAGAGACACTGTGTAACTTCTTAATGATGCTGAGATTC 935
QY 991 GTTCATCTATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTTGATTAAGTATGCTGAAGTG 1050
Db 936 GTTCATCTATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTTGATTAAGTATGCTGAAGTG 995
QY 1051 CCGCAATTTGCAACATCTGTCGCCAGAACATAGGAAAAAGAACAGATGGTGTGA 1110
Db 996 CCGCAATTTGCAACATCTGTCGCCAGAACATAGGAAAAAGAACAGATGGTGTGA 1055
QY 1111 CCGCAATTTGAGAGACTCCATGTAAACAATTAATAAGAAATATTTCAATCTCTTGT 1170
Db 1056 CCGCAATTTGAGAGACTCCATGTAAACAATTAATAAGAAATATTTCAATCTCTTGT 1115
QY 1171 GTTCAGAACTGTTAAA 1187
Db 1116 GTTCAGAACTGTTAAA 1132

RESULT 6
US-09-769-159-1
; Sequence 1, Application US/09769159
; Patent No. US20010021509A1
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH M.
; APPLICANT: SATHE, WENDY S.
; APPLICANT: CHAMBERS, JON
; APPLICANT: MOIR, ALISON
; APPLICANT: SEKERES, PHILIP
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
; TITLE OF INVENTION: ANTAGONISTS OF THE HNEA81 RECEPTOR
; FILE REFERENCE: GH-70318-D3

;; CURRENT APPLICATION NUMBER: US/09/769,159
;; CURRENT FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: 08/956,975
;; PRIOR FILING DATE: 1997-10-23
;; PRIOR APPLICATION NUMBER: 09/221,456
;; PRIOR FILING DATE: 1998-12-28
;; PRIOR APPLICATION NUMBER: 09/558,740
;; PRIOR FILING DATE: 2000-04-26
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 1124
;; TYPE: DNA
;; ORGANISM: HOMO SAPIENS
US-09-769-159-1
Query Match 25.7%; Score 305.6; DB 10; Length 1124;
Best Local Similarity 57.7%; Pred. No. 1,4e-69;
Matches 545; Conservative 0; Mismatches 399; Indels 0; Gaps 0;
QY 152 CTGTGACAGAGACTACAAATACCCAGGTCCCTCCAGCTGCTACATGTCCTG 211
Db 137 CGGTGCCAGAGACACTCGATAGTACAGCTGATATCCAGCCCTCAGACAGTGT 196
QY 212 TTTTGTGACCTTATCAAAATGCGCTGCGATGAGAGATTTCTTCAATCCGAGT 271
Db 197 TTTTGTGACCTTATCAAAATGCGCTGCGATGAGAGATTTCTTCAATCCGAGT 256
QY 272 AAATCAACTTATATTTTCTTAAGACAGATCTATTTGATTTCTATATCTG 331
Db 257 TCTTCACCTTCATCAATCTACCTCAAAACACTTGGGCGACTGATTAATGACATC 316
QY 332 ACTTTCATTAATTTCTAGTATGATGCCAACTGGGAACAGACACAGCACTGAACTTT 391
Db 317 ATGCTTCCTTCAAAATCTCTGACCTACACCTGACCTGACCTGACCTGACCTTT 376
QY 392 GTGTGCAAGTACCTCCGTCATATTTTATTTCAATGATTAATGATTTTCAATCTG 451
Db 377 GTGTGCAAGTACCTCCGTCATATTTTATTTATGACACATGATTAATGATTTTCAATCTG 436
QY 452 GGACGATTAATCAATGCTGCTACCAAGAACACAGCCATTTTAAACATCCACCC 511
Db 437 GGACGATTAATCAATGCTGCTACCAAGAACACAGCCATTTTAAACATCCACCC 496
QY 512 AAAATCTCTTGGGGGCTAAGATCTCTCTGTATCATCTGGGATTCATGCTTACTG 571
Db 497 AAAATCTCTTGGGGGCTAAGATCTCTCTGTATCATCTGGGATTCATGCTTACTG 556
QY 572 TCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATGGAAGAAATG 631
Db 557 TCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATGGAAGAAATG 616
QY 632 TCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATGGAAGAAATG 691
Db 617 TCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATGGAAGAAATG 676
QY 692 GTCAATTTCTGATTAATTTCTTAATGATTTATTTATTTATTTATTTATTTATTTATTT 751
Db 677 TTTATTTCTGACTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 736
QY 752 CTGTACCGGTCATGATGAGAGAGAGGGGTGATGTAAGTCCCGAGAAAAGGTGAC 811
Db 737 GTATATGATTTCTTAATGATTTCCAAAAGTGAAGACAGAAAACAAACAAAGCTGGA 796
QY 812 GTCAAGATTTCTATTAATGATTTGATTTCTTTATTTGTTTCTTCTTCTTCTTCTTCTT 871
Db 797 GGCAGAGATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
QY 872 CGAATTCCTTACACCTGAGCAACCCGGGATGCTTTGATCTGACCTGCTGAAATACT 931
Db 857 AGATTTTCATATCTACAGAGCAACCAAGATTAAGACTGATGAGACTGCAAAATCA 916
QY 932 CTGTCTATGTAAGAGAGACACTGCTGTGTTAACTTCTTAATGATGCTGATCCG 991

```
Db 917 CTTTATTTGCTAAAGAAACACTCTCTTTTGGCAGCAACTAATTTGTATGATCCC 976
Qy 992 TTTCATTTATTTTCCCTTGGCAAGTCTTCAGAAATTCCTTGATTAAGTATGATGTC 1051
Db 977 TTAATATACATATTTCTTATGTAATAAAATTCACAGAAAGTACCATGTATGCAAGGAGA 1036
Qy 1052 CCCAATTCGCAACATCTGTCTCCAGGACATAGGAAAAAGA 1095
Db 1037 AAGACACAGACATCAAGCAAGAAATCATAGCAGTCAGACAGA 1080

RESULT 7
US-09-848-889-2
; Sequence 2, Application US/09848889
; Patent No. US2002002555A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Cheng, Muzong
; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
; FILE REFERENCE: PC-0042 CIP
; CURRENT APPLICATION NUMBER: US-09/848,889
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 568987CBI
US-09-848-889-2

Query Match 25.5%; Score 302.4; DB 10; Length 1488;
Best Local Similarity 57.5%; Pred. No. 1e-68;
Matches 543; Conservative 0; Mismatches 401; Indels 0; Gaps 0;

Qy 152 CTGTGACGACGAGACTACAAATACCCAGTCTCTTCCACTGCTACACTGTCTG 211
Db 131 CGGTGCGCCAGAGACACTGGATAGTACAGCTGTATTCACACCCCTACACAGTGTT 190
Qy 212 TTTTGTGTGGACTATTCACAAATGGCTGGCGATAGAGATTTCTTCAATCCGAGT 271
Db 191 TTCTTACCGGATCTCTGTGTAATCTTGGCTGTGGGTGTGTTCACATCCCACG 250
Qy 272 AATCAACTTTATTTTCTTGAAGAACAGACATTTTGTGATCTTCATGATTCG 331
Db 251 TCTCCACCTTCATCTACTCTCAAAACACTTTGTGGCGCATTTGATATGACACTC 310
Qy 332 ACTTTTCATCAAAATTTCTTAGTATGCCAACTGGGAACAGACACTGAGAACTTT 391
Db 311 ATGCTTCCTTTCAAAATTCCTCTGTGACTCAGACCTGGCAACCCGGAGCTCAGAGCTTT 370
Qy 392 GTGTGCAAGTATCTCCCTCATATTTTATTTACAAATGTATATCATGATTTCCG 451
Db 371 GTGTGCGTGTCTCTGGTGTATTTTATGAGACCATGTATGTGGGCTGCTGTGTTA 430
Qy 452 GACATATATATCATGATGCTACGACAGAGACGACAGGCAATTTAAACATCCAAACCC 511
Db 431 GGGCTATATGCTTTGACAGATTTCCATAGATCATGACCTTTTGAAGAAATATTTTCTA 490
Qy 512 AAAATCTCTTGGGGCTAAGATTTCTCTGTGTCTATCTGGGCATTTGTTCTTACT 571
Db 491 AAAAAACGTGTTTGGCAAAAAGCGTCTCAATCTTCATCTGTGTCTTTTGTCTTCTC 550
Qy 572 TCTTTGCTTAACATGATTTGTGACCAAGCGACGACGAGACAAAGATGTGAAGAAATGC 631
Db 551 TCCCTGCAATATGATCTTGTGAGCAACAGGAAGCAACCATCTGCTGTGAAAAAGTGT 610
Qy 632 TCTTTCCTTAATCAAGAGTGTGCTAGTCTGGCATGAATATGATTAATCAATCTCTCA 691
```

```
Db 611 GCTTCTTAAAGGGGCTCTGGGGCTGAATGGCATCAATGGTAAATACATATGCCAG 670
Qy 692 GTCAATTTCTGATTAATTTCTTAATTTGATATGTAATGTAACACTTACAAAGAA 751
Db 671 TTTATTTTGGAGCTGTATTAATCTTAATGCTGTGTATTTATGGTTATGCAAAAAA 730
Qy 752 CTGTACCGGTATCTTAAGAACGAGGGGTATGTTAAAGTCCCAAGAAAAAGTGAAC 811
Db 731 GTATATGATTTCTTATGAAGAGTCCAAATGTAGAGACGAAAAACAAAAAGCTGGAA 790
Qy 812 GTCAAGTTTCTATTCATCTGCTATCTTATTTGTTTGTGCTTCCATTTTGGC 871
Db 791 GGCAAAGTATTTGTTTGTGCTGCTCTCTTCTTGTGTGTGTTGCTCATTTTCTTGGC 850
Qy 872 CGAATTCCTTACACCTCGAGCCAAACCCGGAGTGTGTTGACTGCTGCTGAATATCT 931
Db 851 AGAGTTCATATCTACAGTCAACCAACAAATACAGTACGTGATGCTGCAAAATCA 910
Qy 932 CTGTCTATGTGAAGAGACACTGTGTGTTACTCTTAAATGCAATGCTGTGATCCG 991
Db 911 CTGTATTTGCTAAAGAAACACTCTTTTGGGACGACACTAACATTTGTATGATCCC 970
Qy 992 TTTCATTTTTCCTTGGCAAGTCTCCAGAAATTCCTTGATAGTATGCTGAGATGC 1051
Db 971 TTAATATCATATTTCTTATGTAATAAAATTCACGAAAGCTCCATCATGTATGCAAGGAGA 1030
Qy 1052 CCCAATTCGCAACATCTGTCTCCAGCAATAGGAAAAAGA 1095
Db 1031 AAGACACAGACATCAAGCAAGAAATCATAGCAGTCAGACAGA 1074

RESULT 8
US-09-919-497-27
; Sequence 27, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIORITY FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 2416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-27

Query Match 21.3%; Score 252.4; DB 10; Length 2416;
Best Local Similarity 55.5%; Pred. No. 1.e-55;
Matches 487; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

Qy 155 TGCACGAGACTACAAATACCCAGTCTCTTCCACTGCTACACTGTCTGTTT 214
Db 256 TCTCTCAAGACCTCTGTGACTCAGAGATCATCTCTGTGTGATGTATGCTTC 315
Qy 215 TTTGTGACTATACAAATAGCGCTGGGATGAGATTTTCTTCAATCCGAGTAAA 274
Db 316 ATTTGGGGATCTTACTCATAGAGTGTGAGATGATTTCTTTTACGTGCCAGCTCT 375
Qy 275 TCAAACTTATTTTCTTGAAGAACAGATCTTGTGATCTCTCTCATGATTTGACT 334
Db 376 AAGAGTTTATCTATCTATCTCAAGAAATGTTTATTTGCTGACTTTGTGATGACCTGACT 435
Qy 335 TTTTCATTCAAATCTTAGTATGCGCAACGTGGAAACAGAACCACTGAGAACTTTGTG 394
Db 436 TTTCTTCAAGATCCTGTGTGACGACGCTTGGTCCCTGGCAGGTGAAGCTGTTTGG 495
Qy 395 TGTCAAGTATCTCGTCTATTTTATTTTACAAATGTATATAGTATTTTCAATCTCTGGA 454
```

Db 496 TGCAGGGCTCTGCCGCTCTTTCTACGTCAACATGTCGTCAGCATTTGTCTTTGGG 555
Qy 455 CTGATTAATCATGATCGCTTACAGAACACACAGGCCATTTAAACATCCACCCAAA 514
Db 556 CTCATACAGCTTTGACAGATATTAATAATTTGAAAGCCTTTTGGACTCTTTATCCAG 615
Qy 515 AATCTCTTGGGGGCTAGATTTCTCTGTGTGTATCTGGCATTCGATCTTTACTCTCT 574
Db 616 TCACTGAGATTACAGCAAACTTCTGTCTCATGTAGATGAGTCATCTCTCTCTCTCT 675
Qy 575 TTGCTACATGATTTCTGACCAACAGAGCCGAGACAGAAATGTAAGAAATGGCT 634
Db 676 GTTCCAAATATTAATTCACACACAGAGCTTTAGGGGTTTACCAAAATTAATGTA 735
Qy 635 TTCCCTTAATCAGAGTTGGTGTCTGTGTCGATGAAATAGTAATTAATCTGTCAAGTC 694
Db 736 GAATGAAAGATGAACTGGAGCGGAGTGGCAAAAGCATCAACTACATCTGTGGCC 795
Qy 695 ATTTCTGAGATTAATTTCTTAATTTGTAATGTATGATTAACATCATTTACAAAGACTG 754
Db 796 ATCTCTGAGATTTGTTCTTTCTTTGTAATCGTTTCTATCTGATCACAAGAAATC 855
Qy 755 TACGGGTATACGTAAGAACGAGGGGTAGTAAAGTCCAGGAAAGGTGAAGTC 814
Db 856 TTTAAGTCCACCTTAAGTCAAGTCGAAATTCACCTTCGTCAAAAGAAATCTAGCCGC 915
Qy 815 AAGTTTCAATTAATCATGCTGTATTTCTTAATTTGTTTGTCTTCCATTTGCCCGA 874
Db 916 AACAATTCAGACATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 975
Qy 875 ATTCCTTACACCCCTGAGCCAAACCCGGGATGCTTTGACTGACATGCTGAAATATCTGTG 934
Db 976 ATCCCTTACAAAGAGTCAAGACGAAACGATCATTAACAGCTCCAGTCAAAAGAAATCTTG 1035
Qy 935 TTCAATGGAAGAGAGACACTGTGTTAATCTTCAATTAATGATCGATGATCCGTC 994
Db 1036 CGGATATGAAAGAAATTCACCTGCTACTATCTGCTGCAAAATGTAAGTGTGGACCTTAT 1095
Qy 995 ATCTATTTTCTCTTGCAGTCTCTTCAAGAAATTCCT 1032
Db 1096 ATTATTTCTTCTATGCGACGCTTATGGAATCTT 1133

RESULT 9
US-09-826-508-19
Sequence 19, Application us/09826508
Patent No. US20010025099A1
GENERAL INFORMATION:
APPLICANT: Nabil Elshourbagy
APPLICANT: Lisa Vawter
TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
FILE REFERENCE: GP-70744USB
CURRENT APPLICATION NUMBER: US/09/826,508
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1017
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-826-508-19

Query Match 19.4%, Score 230.8; DB 10; Length 1017;
Best Local Similarity 53.7%, Pred. No. 2,9e-50;
Matches 478; Conservative 0; Mismatches 412; Indels 0; Gaps 0;

Qy 155 TGCACAGAGACTACAAATACACCAGGTCTCTCCATGCTCTTACACTGTCTGTTT 214
Db 40 TGCTCTGGAACACCTGATCACAAGAGATCATTTCCGTTGTGTACGTAATGTTCTTC 99
Qy 215 TTGTGTGACTTATACAAATGCGCTGGCGATGAGATTTCTTCAATCCGAGCTAAA 274

Db 100 ATCAGGGGCTCTGCCGCTCTTTCTACGTCAACATGTCGTCAGCATTTGTCTTTGGG 159
Qy 275 TCAACCTTATTAATTTTCTTAAGAACACAGTCAATTTCTGATCTCTCATATTTGACT 334
Db 160 AAGATTTATCATCTATCTAAGACATAGTGGTGTGACTTTCTTATGGCTGACT 219
Qy 335 TTTCATTAAGAAATTTCTAGTATGCCAACTGGGACAGACACATGAGAACTTTTGTG 394
Db 220 TTCCCTTCAAAAGTCTTGTGACTCAGAGCCCTGGCCCTGGAGTGAATGTTGTGTG 279
Qy 395 TGTCAAGTACTCTCCGATATTTTATTTACAAATGTAATATACATTTTCTCTGGGA 454
Db 280 TGCAGGGCTCTGCCGCTCTTTCTATGTAATGTAATGTAATGTAATGTAATGTAATG 339
Qy 455 CTGATTAATCATGATCGCTTACAGAACACACAGGCCATTTAAACATCCACCCAAA 514
Db 340 CTCATACAGCTTTGACAGATTAATTAATTTGTAAGCCCTTCTGACCTCTATTTGTGAG 399
Qy 515 AATCTCTTGGGGGCTAGATTTCTCTGTGTGTATCTGTCATCTGGGCAATTCATCTCTCT 574
Db 400 TCGGGAATATAGCAACGCTCTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
Qy 575 TTGCTACATGATTTCTGACCAACAGAGCCGAGACAGACAAAGATGTAAGAAATGCTCT 634
Db 460 GTCCAAACATCATCTCTACAAACACAGGGGTCAAGGAGGTACAGAAATACAGTCAATG 519
Qy 635 TTCCCTTAATCAGAGTTGGTGTCTGTGTCGATGAAATAGTAATTAATCTGTCAAGTC 694
Db 520 GAGCTCAAAACAGAGCTGGGGGAGAGTGGCAAAAGGCTTACATATTTCTGTGTGT 579
Qy 695 ATTTCTGAGATTAATTTCTTAATTTGTAATGTATGATTAACATCATTTACAAAGACTG 754
Db 580 ATCTCTGAGTGTGTTCTTCTGTCTGTAATGCTTCTACAGGCGCATCAGAGAAATG 639
Qy 755 TACGGGTATACGTAAGAACGAGGGGTAGTAAAGTCCAGGAAAGGTGAAGTC 814
Db 640 TTCAAGTCTACCTTCAAGTCCAGGAATTTCCACCTCCGTCGTAAGGAAAGTCAAGCCGC 699
Qy 815 AAGTTTCAATTAATCATGCTGTATTTCTTAATTTGTTTGTCTTCCATTTGCCCGA 874
Db 700 AATATCTTACAGATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
Qy 875 ATTCCTTACACCCCTGAGCCAAACCCGGGATGCTTTGACTGACATGCTGAAATATCTGTG 934
Db 760 ATCCCTTACAAAGAGTCAAGACGAAAGTCACTACAGCTGCGGACCAAGAGACCTGT 819
Qy 935 TTCAATGGAAGAGAGACACTGTGTTAATCTTCAATTAATGATGCTGTGATCCGTC 994
Db 820 CTCTATGGAAGAAATTCACCTGCTACTCTGCTGCAATGTTGTCTGTGACCCCATTT 879
Qy 995 ATCTATTTTCTCTTGCAGTCTCTTCAAGAAATTCCTGATAAGTATGCT 1044
Db 880 ATTATTTCTTCTATGCGACGATTTAGAGAAAGTCTTAATTAAGAAATTT 929

RESULT 10
US-09-826-508-9
Sequence 9, Application us/09826508
Patent No. US20010025099A1
GENERAL INFORMATION:
APPLICANT: Nabil Elshourbagy
APPLICANT: Lisa Vawter
TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
FILE REFERENCE: GP-70744USB
CURRENT APPLICATION NUMBER: US/09/826,508
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 1077
TYPE: DNA
ORGANISM: HOMO SAPIENS

US-09-826-508-9

Query Match 17.7%; Score 209.8; DB 10; Length 1077;
Best Local Similarity 53.4%; Pred. No. 8.1e-45;
Matches 464; Conservative 0; Mismatches 402; Indels 3; Gaps 1;

157 CACGAGAGCTACAAATCACCAGGCTCTTCCACTGCTCTACACTGCTGCTTTT 216
102 CACCTTCACAAATGATGACACAAATGCTGCGCTTATCTCATATATATGTT 161
217 TGTGACCTTATCACAATGGCCCTGGCAGATGAGATTTCTTCAATCCGATAATC 276
162 GCGAAGCATCTGTGTAATGTTAGAGATGATGATCTCTTCCACATTTAGATTAAC 221
277 AACTTTATTTATTTCTTAAAGACACAGTCAATTTGATCTCTCAATGATCTGACTT 336
222 CAGCTTCATATTCATCTCAAAAACATAGTGTGACAGCTCTCATATGACGCTCAT 281
337 TCCATTCAAATTTCTGATGACCAACCTGGGACAGAACGACTGAACTTTGTGTG 396
282 TCCATTTGCAATGATCATATGACAGATTTGACCTTGTACTTCAAGTTTATCTGTG 341
397 TCAAGTACCTCCGCTATTTTATTTTCAACATGATATGATATTTCAATTTCTGGAGCT 456
342 CAGATACACTTCAATTTGTTTATGCAAAATGATATCTTCCATCTGTTCTTCTG 401
457 GATTAATGATGATGCTTACAGAGAACCCAGGCTTTTAAACATCCAAACCCCAAAA 516
402 GATTAAGATGATGATGCTTATGAGAGTGTGCAAGCTTTGGGAGCTCTGATGTACAG 461
517 TCTCTTGGGGGCTAGATCTCTGTTGTCATCTGGGCTATGCTTCTTACTCTTCT 576
462 CATACCTTCACGAAGGTTTATCTGTTGTTGTTGGGATCAATGCTTTGCTCTT 521
577 GCTTAACATGATTTCTGACCAACAGGCGGAGAGACAAATGATGAGAAATGCTT 636
522 GCGAAGCATCTGCTGTAATGTTAGAGATGCTGACCAAGAGACAAATGATGATGCTT 581
637 CTTAATGATGATGCTTACAGAGAACCCAGGCTTTTAAACATCCAAACCCCAAAA 516
582 ACTTAAGATGATGCTTATGAGAGTGTGCAAGCTTTGGGAGCTCTGATGTACAG 461
697 TTTCTGATTAATTTCTTATTTGTTATGTTATGATCTATTCACAAAGACTGTA 756
642 GTTTGAGCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 701
757 CCGTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
702 CAATTCAGAGGCAATTCATTAAGT---CAGTCAAGCCGAAAGGAAACATTAACCGAG 758
817 AGTTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
759 CATCAGGTTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 818
877 TCTTACACCTGACGCAAAACCGGGATGTTGATCTGACAGCTGTAATTAATCTGCT 936
819 TCTTTTACTTTTATGCTATGACAGAGCTTTTATGATGATGATGATGATGATGATGAT 878
937 CTATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 996
879 TTAAGTGAAGAAATTAACATTTTCTGCTGCTGATGATGATGATGATGATGATGATGAT 938
997 CTATTTTCTTTTCTTTGCAAGTCTTCAAGAA 1025
939 TTACTTTTCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967

RESULT 11
US-09-974-298-85
; Sequence 85, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei

TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIORITY APPLICATION NUMBER: 60/238,331
; PRIORITY FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 85
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1650519CB1
US-09-974-298-85

Query Match 17.5%; Score 207.4; DB 9; Length 1444;
Best Local Similarity 53.0%; Pred. No. 3.8e-44;
Matches 467; Conservative 0; Mismatches 411; Indels 3; Gaps 1;

157 CACGAGAGCTACAAATCACCAGGCTCTTCCACTGCTCTACACTGCTGCTTTT 216
420 CACCTTCACAAATGATGACACAAATGCTGCGCTTATCTCATATATATGTT 479
217 TGTGACCTTATCACAATGGCCCTGGCAGATGAGATTTCTTCAATCCGAGTAATC 276
480 GCGAAGCATCTGTGTAATGTTAGAGATGATGATCTCTTCCACATTTAGATTAAC 539
277 AACTTTATTTATTTCTTAAAGACACAGTCAATTTGATCTCTCAATGATCTGACTT 336
540 CACCTTCATATTTCTATCTCAAAAACATAGTGTGCAAGCTTATATGACGCTGACAT 599
337 TCCATTCAAATTTCTAGATGCTCAAACTGGGACAGACGACACTGAGAACTTTGCTG 396
600 TCCATTTGCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
397 TCAAGTACCTCCGCTATTTTATTTTCAACATGATATGATATTTCAATTTCTGGAGCT 456
660 CAGATACACTTCAAGTTTGTGTTTATGCAAAACATGATATCTTCCATGCTGTTGCT 719
457 GATTAATGATGATGCTTACAGAGAACCCAGGCTTTTAAACATCCAAACCCCAAAA 516
720 GATTAAGATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
517 TCTTGGGGGCTAAGATCTCTGTTGTCATCTGGGCTATGCTTCTTACTCTCTT 576
780 CATACCTTCAGAGATTTTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 839
577 GCTTAACATGATTTCTGACCAACAGGCGGAGAGACAAAGATGTAAGAAATGCTT 636
840 GCGAAGCATCTGCTGTAATGTTAGAGATGCTGACCAAGAGAGCAATATCCATGATGCT 899
637 CTTAATGATGATGCTTACAGAGAACCCAGGCTTTTAAACATCCAAACCCCAAAA 516
900 ACTTAAGATGATGCTTATGAGAGTGTGCAAGCTTTGGGAGCTCTGATGTACAG 461
697 TTTCTGATTAATTTCTTATTTGTTATGTTATGATCTATTCACAAAGACTGTA 756
960 GTTTGAGCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 701
757 CCGTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
1020 CAATTCAGAGGCAATTCATTAAGT---CAGTCAAGCCGAAAGGAAACATTAACCGAG 1076
817 AGTTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
1077 CATCAGGTTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 818
877 TCTTACACCTGACGCAAAACCGGGATGTTGATCTGACAGCTGTAATTAATCTGCT 936
1137 GCCTTCTACTTTATGCTATGACAGAGCTTTTATGATGATGATGATGATGATGATGATGAT 1196

QY 937 CTATGTAAGAGGACGCTGTGGCTAACTTCTTAAAGCATGGCTGGATCCCTCAT 996
 Db 1197 TTACTGCAGCAACAAATTCACCTTTCTGTGTGCTGTGAATGTTGGCTGCATCAATTAAT 125
 QY 997 CTATTTTTCCTTTCGACAGTCCTCAGAAATTCCTTGATAA 1037
 Db 1257 TTACTTTTCATGTGTAGTCATTTTCAGATGGCTGTGA 1297

RESULT 12

```

US-09-919-172-21
Sequence 21, Application US/09919172
Patent No. US20020119463A1
GENERAL INFORMATION:
APPLICANT: Faris, Mary
APPLICANT: Turner, Christopher M.
TITLE OF INVENTION: PROSTATE CANCER MARKERS
FILE REFERENCE: PA-0036 US
CURRENT APPLICATION NUMBER: US/09/919,172
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/222,469
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PERL Program
SEQ ID NO 21
LENGTH: 1444
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020119463A1 1650519CBA1
US-09-919-172-21

```

[illegible]

QY	697	TTTGTGATTAATTTCTTAA	TGTATTTGTATGTATATACACTCATATACAAAAGACTGA	756
Db	960	GTTTGTGCGCGGTGCTGGTGTATCTGTATCGAGATGATACATAGCCATATCCAGTACATCCA	1019	
QY	757	CCGTCATACGTAAAGAACGAGGGGTGTAGGTAAAGTCCCCAGCAAAAAGTGAACGTCAA	816	
Db	1020	CAAAATCCAGCAGCAATCTTAAGT---	CAGTCAGACCGCAAAAGCGAAAACATATACCAAGC	1076
QY	817	AGTTTTCATTATCATGTCACCTGTATCTTATATTTGTTTGTTCCTTCCATATTTGCCCGAAT	876	
Db	1077	CATCAGGGTGTGTGTGTGCTGTATTTTACCTGCTTTTACACATATACACTTGTGCAGAAT	1136	
QY	877	TCCTTACACCGCTGAGCCAAACCCGGGATGTCTTGTACTGCACCTGCTGAAAATACTCTGT	936	
Db	1137	GCCTTCTACGTTTATGATCAGTTTACAGAGGCTTTTATAGATGAAATCGACAACAAAATCCTATA	1196	
QY	937	CTATGTCAAAAGAGACACTGTGTGTAACTTCCTTAAATGATGATCCCTGATCGCTTCAT	996	
Db	1197	TTACTGCAAAAGAAATATACACTTTTCTGTGCTGTGGTATATGTTTCCCTGGATCCAAATAT	1256	
QY	997	CTATTTTTCCTTTGCAAGCTTCCTCAGAAATCCTTGATAA	1037	
Db	1257	TTACTTTTCACTGTGTAGCTCATTTTCCAAAGATGGCTGTCA	1297	

RESULT 13

```

US-09-848-889-9
Sequence 9, Application US/09848889
Patent No. US2002002555A1
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Gueglar, Karl J.
APPLICANT: Cheng, Muzong
TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
FILE REFERENCE: PG-0042 CIP
CURRENT APPLICATION NUMBER: US/09/848, 889
CURRENT FILING DATE: 2001-05-03
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 592
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO. US2002002555A1 6867412HT
US-09-848-889-9

```

	Query Match	Similarity	16.6%	Score 197.4;	DB 10;	Length 592;
	Best Local	Similarity	59.3%;	Pred. No. 9.9e-42;		
	Matches	Conservative	59.3%;	Mismatches 231;	Indels	Gaps
Oy	147	CCAGCTGNGCACCGACGACTCAAAATTCACCAGGTCTCTTCCACGCTCTACACG	206			
Dd	13	CTAACGGTGCCCCAGAGACACTCGGAATGTCACGCTGTATTCCAGGCCCTC	72			
Oy	207	TCTCTTTTTTTGGTTGGACTTATCACAAATGGCCCTGGCGATGAGATTTTCTTCAATCC	266			
Dd	73	TGCTTTCTCTGACGGGCATCTCGTGAAATACTTTGCTGTGGGTGTTTGTTCACATCC	132			
Oy	267	GGAGTAAATCAAACCTTATTTATTTTCTTAAGACACAGCTCATTTCTGATCTTCATGA	326			
Dd	133	CCAGCTCTCTCCACCTTTCATCATCTACCTTCATAAACAACCTTGGTGCGCCACTTGATTAATCA	192			
Oy	327	TTTCACACTTTCOCATTCAAAAATTTCTAGTAGAGCCCAACTGGGAACAGACACATGAGAA	386			
Dd	193	CACGTATGCTTCTCTTCAAAATCTCTCTGACATCAACCTGGACCCGTGGACGCTCAGAG	252			
Oy	387	CTTTTGTGTGTCATGATTAACCTCCGTCATCTTTTATTTTCAATGTATATACATTTTTCAT	446			
Dd	253	CTTTTGTGTGTCGTTTTTCTTCGGTGATATTTTATGAGACCATGTATGTGGGCATCTGTC	312			
Oy	447	TCTTGGAAGTAACTATGATGCTACACGAGACCAACCGGCCATTTTAAACATCCA	506			

Db 313 TGTATAGGCTATACGCTTTGACAGATTCCTCAAGATCAGACCTTTGAGAAATATTTT 372
QY 507 ACCCAAAATCTCTG6GGGCTAAGATTCTCTGTGTGATCGGGCATTCATGTTCT 566
Db 373 TTTTAAAAAAGCTGTTTTTGCAAAAAGGCTTCATCTCATCTGTTCTTTTGTCT 432
QY 567 TACTCTCTTGGCTTAACATGATTTCTGACACAGCGCAGGAGACAGATGTAAGA 626
Db 433 TCATCTCCCTCCCAATATGATCTTAGCAACAAGGACACACATCTCTGTGAAA 492
QY 627 AATGCTCTTTCCTTAATATGAGATGCTGCTAGCTGAGTAAATAGTAAATTAATCT 686
Db 493 AGTGTCTCTTCTTAAAGGGGCTCTG6GGCTGAATG6CATCAATGTAATACATAT 552
QY 687 GTCAAGTCATTTTCTGATTAATTTCT 713
Db 553 GCCAGTTTATTTCTGACTGTTTAT 579

RESULT 14

US-10-025-335-2

; Sequence 2, Application US/10025335

; Patent No. US20020127648A1

; GENERAL INFORMATION:

; APPLICANT: Coleman, Roger

; Bandman, Olga

; TITLE OF INVENTION: NOVEL HUMAN CSA-LIKE RECEPTOR

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/025,335

; FILING DATE: 18-Dec-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/791,974

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0198 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1257 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: Consensus

; CLONE: 346874

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-025-335-2

Query Match

Best Local Similarity 11.1%; Score 131.8; DB 12; Length 1257;

Matches 429; Conservative 1; Mismatches 443; Indels 12; Gaps 2;

QY 142 GAACACCACTGTCGACACGAGACTACAAATACACCCAGGTCCTTCCCACTGCTCA 201

Db 208 GACAAACAGTTCGTTCTTCCGACGTTTATAAGATCTGGACCATCACTGATTTT 267
QY 202 CACTGTCCTGTTTTTTGGTGAATTCACAAATGGCGGCGATGAGATTTCTTTCA 261
Db 268 TTTATTAGTTTCTCTGTTGGAAATTTATGGAATGTTTTGCAACCTGGGCTTTTATCA 327
QY 262 AATCCGAGTAATCAAACT--TTATTATTTTCTTAAGAACACAGTCAATTTGATCT 318
Db 328 GAAGATACGAATCACAGGCTGTGAGCATCTACTTAATTAATTGCTTACACGATTT 387
QY 319 TCTCATGATTCATCTTTTCCATTTCAAAATCTTATGATGCCAACTGGGAACAGACC 378
Db 388 CCGCTTACTCTGGCAATTTACCAAGTGAATAATGTTGTTGATCTGGGCTGACCTTGGA 447
QY 379 ACTGAAACCTTTGTTGTCAGTTCCTCCGATATTTTATTTTACAGATATATCAG 438
Db 448 RCTGAAGATATTCACATGCCAGTACAGGCTCCCTCATCTATATCATATGATTTATC 507
QY 439 TATTCAATTCCTGGGACTGATTAATCATGATGCTACAGAACACCAAGCCATTTAA 498
Db 508 AATTATCTTCTTATGATTTGTCAGATTTGACGCTGCTTACAGCTGACACAGCTGAA 567
QY 499 AACTCCAAACCCAAATCTCTGGGGCTTAAGATTTCTCTGTTGTCATCTGGCATT 558
Db 568 GATCTACCGAATACAGAACCCGCTTTCACAAATGATATCAACGCTGTGCTTAAT 627
QY 559 CATGTCCTTACTCTCTTGGCTTACATGATTTGACCAACAGGACCGGAGAACAA 618
Db 628 GGTCTTCTTATATGCTGCAAAATATGATTTCCATCAAGACATCAAGAAAGTC 687
QY 619 TGTGAAGAAAGCTCTTCTTAAATCAGAGTTCGCTAGCTGAGTGAATAGTAA 678
Db 688 AATGTGGGTGTATGAGATTTAAAGAAATTTGGAATAATTTGCTGCTGACAA 747
QY 679 TTACATCTGTCAGATTTTTCGATTAATTTCTTAATTTGTTATGTTATATACCT 738
Db 748 TTTCAATATGTCAGATTTTAAATTTCTCAGGCAATTTTAAATTTCAATTTGCT 807
QY 739 CATTAACAAAGAACTGACGCTATGATGTAAGAACAGAGGGGTGTGTAAGTCCCA 798
Db 808 TGTAAATTCGACAGCTTACAG-----AAACAAAGATATGAATAATTCACCAAT 858
QY 799 GAAAAAGTGAACGTCGAAGTTTTCATTAATTCATTTGCTGATTTTATTTGTTGTC 858
Db 859 GAAAAAGCTTCATCAATCTTTTAAATTTCTCAGGCAATTTTAAATTTCAATTTGCT 918
QY 859 TTTCCATTTTGGCCGATTCCTTACACCTGAGCCAAACCCGGAGTCTTTGACGAC 918
Db 919 TTACCAATTTGTCGATCCGATTCCTTACCTCAGCCAGACAGAGATCATATGATGCTC 978
QY 919 TGTGAAATTAATCTGTTCTATGTAAGAGAGACACTGTGTTTAACTTTAAATG 978
Db 979 AACCAAGATTTCACTTTTAAAGCCAAAGAGGCTACAGCTGCTGCTGCTGCAACT 1038
QY 979 ATGCTGCAATCCGTTATCTATTTTCTTCTTGAAGTCTTCAAG 1023
Db 1039 GTGCTTGTATCTATCTGATCATCACTTCAAAAGATTCG 1083

RESULT 15

US-09-848-889-10

; Sequence 10, Application US/0984889

; Patent No. US2002002555A1

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; Applicant: Guegler, Karl J.

; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER

; FILE REFERENCE: PC-0042 CIP

; CURRENT APPLICATION NUMBER: US/09/848,889

; FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 12

```
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 7979275H1
US-09-848-889-10
```

```
Query Match      10.7%; Score 126.6; DB 10; Length 518;
Best Local Similarity 57.8%; Pred. No. 1.9e-23;
Matches 225; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
```

```
QY 152 CTGTGCACGAGAGACTCAAAATCACCAGTCCCTTCCACGCTCTACACGTCCTG 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 CGGTGCCCGACAGACTCGAATAGTACAGCTGATTCGACGCCCTCTACACAGTGT 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 TTTTGTGTGACTTATCACAATGGCCTGGCGATGAGATTCTTCAAAATCCGAGT 271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 TTCTTGACCGGCAATCCCTGCTGAATACTTGTGCTGTGGGTGTTGTTCACATCC 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 AAATCAACCTTTATTTTCTTAAGACAGTCATTTCTGATCTTCATGATTCGTG 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 TCCTCCACCTTCATCACTACCTCAAAACACTTGGTGCCGACTTGATATATGACATC 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 ACTTTCCATTCAAAATTTCTTAGTATGCCAACTGGGAACAGACCACTGAGAACTTT 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 ATGCTTCTTTCAAAATCCTCTCTGACTCACACCTGGCACCTGGCAGCTCAGAGCTTT 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 GTGTGTCAGTTACCTCCGTCAATATTTTATTTACAAATGTATATCAGTATTTCAATCCTG 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 GTGTGTCGTTTCTCTGCGTATATTTATGAGACATGTATGCGCATCGTGTGTGA 427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 452 GGACTGATTAATCGATGCTACGCTACAGACAGACCGGCAATTAACATCCAAACCC 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 GGGCTCATAGCCTTTGACAGATTCCTCAAGATCTCAGACCTTTGAGAAAATTTTCTA 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 512 AAAAATCTTTGGGGGCTAAGATTCCTCTC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 AAAAATCTTTTGGCAAAACGGTCTC 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 11, 2002, 09:13:09
Job time : 120 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 03:09:10 ; Search time 2197 Seconds

(without alignments)
8750.133 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187
Sequence: 1 ggtaccggggagatctctag.....tgytccagaactcgttaa 1187Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_estnum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	732.2	61.7	1541	AK013804	AK013804 Mus muscu
2	732.2	61.7	2103	AK014807	AK014807 Mus muscu
3	638.2	53.8	937	AK014807	AK014807 Mus muscu
4	623.8	52.6	680	AK014807	AK014807 Mus muscu
5	616.8	52.0	1043	AK014807	AK014807 Mus muscu
6	600	50.5	738	AK014807	AK014807 Mus muscu

C	7	599.6	50.5	609	17	AK0581896
	8	577.2	48.6	971	14	B0893760
	9	507.8	42.8	511	14	BM987216
	10	505.4	42.6	706	10	AV718096
	11	449.4	37.9	674	10	BB624828
	12	442.4	37.3	1681	14	BM926255
	13	424.2	35.7	877	14	BM919208
	14	407.6	34.3	542	9	AA274112
	15	366.8	30.9	915	12	BF677440
	16	334.8	28.2	476	10	AM045980
	17	307.2	25.9	1312	11	AF345565
	18	249	21.0	265	17	AK0581861
	19	246.2	20.7	2515	11	AK008013
	20	240.8	20.3	257	14	BM713326
	21	240.8	20.3	262	14	BM674673
	22	212.6	17.9	775	11	BC029363
	23	208.2	17.5	1199	11	AF345566
	24	203	17.1	729	13	BI489709
	25	190.8	16.1	800	12	BG621718
	26	180.2	15.2	961	13	BI103980
	27	170.8	14.4	668	9	AU168581
	28	162.8	13.7	851	12	BG677820
	29	154.2	13.0	455	9	AI090920
	30	153.4	12.9	211	10	BE687381
	31	151.6	12.8	615	10	BB632376
	32	150.8	12.7	669	10	BB653415
	33	149.6	12.6	622	14	BM935684
	34	142.2	12.0	467	9	AT743546
	35	132	11.1	421	9	AI028704
	36	128.8	10.9	554	17	AQ241005
	37	126.6	10.7	790	12	BG205036
	38	123.8	10.4	418	10	AW295791
	39	123.6	10.4	593	12	BE879604
	40	123	10.4	478	12	BF565992
	41	121.4	10.2	976	14	BO900494
	42	119.8	10.1	813	10	BE563611
	43	114.2	9.6	586	17	BH327804
	44	113.8	9.6	360	9	AI656746
	45	113.6	9.6	519	9	AI050350

ALIGNMENTS

RESULT 1
AK013804
LOCUS
DEFINITION
Mus musculus adult male hippocampus cDNA, RIKEN full-length
RECEIVED, full insert sequence.
ACCESSION
AK013804
VERSION
AK013804.1 GI:12851306
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus
clone:1900079B22.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scleroglossa; Muridae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
99279253
10349636
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159

Db	863	CCAAAGAGGAGCCAAAAGATTAAGAGAGTAAACAAAATGTTCTTTCTTAAGTCAGAGTTTG	922
OY	654	GTCAGTCGTGGCATGAATAGTAATTAATACATCTGTCAAGTCATTTTCTGGATTAATTTCT	713
Db	923	GTCATGATTTGGCAGAAATAGTCATTAATTAATCATCTGCCAAGTCATTTTCTGGATTAATTTT	982
OY	714	TAAATGTTATTTGATGTTATACCTATTAACAAAGAAAGTGAACGGGTCATACGTAAAGA	773
Db	983	TAAATGTCATCGTTTGTTATTAAGCCTCAATTAACAAAGATCCTATGCTTAATGTCAGAA	1042
OY	774	CGAGGGGTGATAGGAAAGTCCCGAGGAAAAAGGTGAACCTCAAGATTTTCATTAATTTG	833
Db	1043	CAGAGGTTTCAGCCAAAGTTCCCAAGAAAAAGGTAAAGCTCAAGGTTTTCATCATCATTG	1102
OY	834	CTGATATCTTTATTTGTTTGTTCCTTCCATTTTGGCCGAATTCCTTACACCCGTGAGCC	893
Db	1103	CTGATATCTTTATTTGCTTTGCTTCCCTCCACATTTTCACGAGATTCCTTACACCCGTGAGCC	1162
OY	894	AAACCCGGGATGCTTTGACATGCACTGCTGAGAAATCTGTGTTCTATGTGAAGAGACA	953
Db	1163	AAATCTCGGGGCGCTTTGGACTGTGAGCTGAGAGAACCCCTGTTCTACAGTGAAGGAGACA	1222
OY	954	CTCTGCTGTTAACTTCCTTAAATGATGATGCTGATCCGTTCAATCTATTTTCCATTTGCA	1013
Db	1223	CCCTATGCTGACGTCACTACAGCCCTGCTTGTATCATTAATCACTATTTTCTTTGCA	1282
OY	1014	AGTCCTTCAGAAATTCCTTTGATAGTATGCTGAAAGTGCCCCCAATTTCTGCACATCTCTGT	1073
Db	1283	AGCTTTTCAGAAATTCCTTTACCAAGCATGCTGAGTGCTTCACAAACCTCT--ACATACAAAT	1339
OY	1074	CCCAAGCAATAGGAAAAAAGAAAGAGATGCTGATACCCAAATGAGAGACTCCAAATGT	1133
Db	1340	CTGGGAGCAAAACAAGAAAGAAAGCAAGAAAGGTGGCAACCAAGCAAGAGACCCCAATGT	1399
OY	1134	AAACAATTAACTAGGAATATTTCAATCTCT	1166
Db	1400	AGA-ACATTAACCAAGGGGCTGCTTTACGTCCTTT	1431

RESULT 2	
AK014807	
LOCUS	AK014807 2103 bp mRNA linear HTC 19-JAN-2002
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921504D23;putative cygic receptor Pzyr, G-protein coupled 12, full insert sequence.
ACCESSION	AK014807
VERSION	AK014807.1 GI:12852867
KEYWORDS	HTC, CAP trapper.
SOURCE	Mus musculus (Strain:C57BL/6J) adult male testis cDNA to mRNA, clone_11b;RIKEN full-length enriched mouse cDNA library clone:4921504D23. Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
AUTHORS	JOURNAL MEDLINE 99279253 PUBMED 10349636
REFERENCE	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	JOURNAL MEDLINE 20499374 PUBMED 11042159
REFERENCE	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasliwalgi,K.,

TITLE	JOURNAL	PUBMED	REFERENCE
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiz, A. and Hayashizaki, Y. RIKEN Integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multiplexillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	20530913	
Kawai, J., Shinaagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		11076861	4

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5 (bases 1 to 2103)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,

TITLE Direct Submission
JOURNAL Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama Kanagawa 230-0045, Japan (E-mail: ygenome-res@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

FEATURES	Location/Qualifiers
source	1. .2103

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:4921504D23"
/db_xref="MGD,MGI:1903300"
/db_xref="taxon:10090"
/clone="4921504D23"
/sex="male"
/tissue_type="testis"
/clone_1ib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..2103
/gene="P2ry12"
211..1254
/gene="P2ry12"
/note="data source:MGD, source key:MGI:1918089,
evidence:ISS
purinergic receptor P2Y, G-protein coupled 12
putative"
/codon_start=1
/protein_id="BAB29561.1"
/db_xref="GI:12852868"
/db_xref="MGD,MGI:1918089"
/translacion="MDVGVNNTSANTFSPGTSLCYRDYKITQVLPPLTYLFEFA
GLTNSLARIFFOIRSKNFIFLKNVISDLMLFEPKILSDATLGGPRLTV
COVTSYTFYFMYSISPLGLITDRYKTRPRKTSPLSLKAKILSYVIMAFEL
ISLPMILITNRPRKDKDYTKCSFLKSEFGLVHWEIVITICQVTFINFLIVICSLI
TKELRYSTVRGSAKVPKRVKVKVFIILAVFLICEVPEFHPARIPYLSQTRAVFDC
SAENITLFFVKESTLMLSLNACLDPIFFIFLCKSPRSLTSMLRCSNSTSTGTRKKK
GOEGEPSEETPM"
2083..2088
/gene="P2ry12"
/note="putative"
2103
polyA_site
/gene="P2ry12"
/note="putative"
2103
polyA_signal
2083..2088
BASE COUNT 677 a 448 c 360 g 618 t
ORIGIN
Query Match 61.7%; Score 732.2; DB 11; Length 2103;
Best Local Similarity 82.2%; Pred. No. 4,8e-149;
Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;
QY 114 CCGTGCACAAATCTCACCTCTGGCCTGGGAACACAGCTGTGCGACCAAGACTACAAA 173
DB 236 CCTCAGCCAAATACCACTCTCCCTGGGACACACCTGTGGCGACAGACTACAAA 295
QY 174 TCACCCAGGCTCTCTCCACAGCTCTACACAGCTGTGTTTGTGGACTTATCAAA 233
DB 296 TCACCCAGGCTCTCTCCACAGCTCTACACAGCTGTGTTTGTGGACTTATCAAA 355
QY 234 ATGGCTGGCGATGAGAGATTTCTTCAATCCGAGATTAATCAACTTATTTATTTTC 293
DB 356 ACAGCTTGGCAATGAGAGATTTCTTCAATCCGAGATTAATCAACTTATTTATTTTC 415
QY 294 TTAAGACAGCTCATTTCTGATCTTCATGATTCGATTCATTCATCAAAATCTTA 353
DB 416 TTAAGACAGCTCATTTCTGATCTTCATGATTCGATTCATTCATCAAAATCTTA 475
QY 354 GTGATGCCAAACTGGGAACAGGACACAGCACTTGTGTGTCAGATTACCTCCGTA 413
DB 476 GTGATGCCAAACTGGGAACAGGACACAGCACTTGTGTGTCAGATTACCTCCGTA 535
QY 414 TATTTTATTTTCAATGATATATACATTTTCTTCCCTGGGACTGATTAATGATCGCT 473
DB 536 CATTTTATTTTCAATGATATATACATTTTCTTCCCTGGGACTGATTAATGATCGCT 595
QY 474 ACCAAGAGCCAGCGCATTTAAACATCCAAACCAAAATCTCTGGGGGCTAAAA 533
DB 596 ACCAAGAGCCAGCGCATTTAAACATCCAGCCCAAGCAATCTCTGGGGGCTAAAA 655
QY 534 TTTCTCTCTGTCATCTGCGGATTCATGTTCTTCTCTTGTGCTTACATGATGATCTGA 593
DB 656 TTTCTCTCTGTCATCTGCGGATTCATGTTCTTCTTATTTTCACTGCTTACATGATTTCTCA 715

```

```

QY 594 CCAACAGGACGCCGAGAGACAAAGAAATGTGAGAAATGCTCTTCTTAAATCAGATTGC 653
DB 716 CCAACAGGAGGCCAAAGAGATTAAGACGTAACAAATGTTCTTCTTAAAGTCAGAGTTTG 775
QY 654 GTCGTATCTGGGATGAATAGTAATATACATCTGTGCAAGTCAATTTTCTGATTAATTTCT 713
DB 776 GTCGTATCTGGGATGAATAGTAATATACATCTGTGCAAGTCAATTTTCTGATTAATTTTCT 835
QY 714 TATTTTATTTTATTTATTTATACATCATTTACAAAGAAATGTGACGGTCAATGATTAAGAA 773
DB 836 TATTTTATTTTATTTATTTATACATCATTTACCAAAAGAAATGTGATTTTATTTATTTTAT 895
QY 774 CGAGGGGTGTAGTAAAGTCCCAAGAAAGAAAGTGAAGCTCAAGTTTTCATTAATCATTTG 833
DB 896 CAAGGGGTGTAGTAAAGTCCCAAGAAAGAAAGTGAAGCTCAAGTTTTCATTAATCATTTG 955
QY 834 CTTGATTTCTTTTATTTTGTGTTTGTCTTCCCTTCCATTTTCCCGAATTTCTTACACCTGAGGC 893
DB 956 CTTGATTTCTTTTATTTTGTGTTTGTCTTCCCTTCCATTTTGCACGGAATTTCCCTACACCTGAGGC 1015
QY 894 AAACCCGGGATGTCTTTGACATGCTGCTGAATAATCTGTTCTATGTGAAGAGAGA 953
DB 1016 AAACCTGGGCGCTCTTTGATCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1075
QY 954 CTCTGTGTTAACTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013
DB 1076 CCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
QY 1014 AGTCTTCAAGAAATCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073
DB 1136 AGTCTTCAAGAAATCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
QY 1074 CCCAGGACAAATAGGAAAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
DB 1193 CTTGGGACAAAGAAAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
QY 1134 AAACAAATTAATAGGAAATTTCAATCTCT 1166
DB 1253 AGA-ACATTAACCAAGGCGCTCTTACGCTT 1284

```

RESULT 3
B1758761
LOCUS
DEFINITION 603023751P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194356 5',
mRNA sequence.
B1758761
ACCESSION B1758761.1 GI:15750339
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 937)
AUTHORS NIH-MGC http://mgi.mgi.nhl.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nsl.nhl.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1486 row: e column: 13
High quality sequence start: 28
High quality sequence stop: 853.
Location/Qualifiers
1..937
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/Note="Organ: brain; Vector: PCMV-SpOrf6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT      273 a      227 c      164 g      272 t      1 others
ORIGIN

```

```

Query Match      53.8%; Score 638.2; DB 13; Length 937;
Best Local Similarity 94.0%; Pred. No. 1.4e-128;
Matches 719; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

```

```

QY 28 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 87
D 174 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 233
QY 88 ATCAGGTACCAAGAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 147
D 234 ATCAGGTACCAAGAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 293
QY 148 CAGTCTGTGACACGAGACTACAAATCACCAGGTCCTCTCCACGCTCTACAGCT 207
D 294 CAGTCTGTGACACGAGACTACAAATCACCAGGTCCTCTCCACGCTCTACAGCT 353
QY 208 CCGTCTTTTGTGGACTATACAAATGGCCTGGCATGAGATTTCTTCAATCCG 267
D 354 CCGTCTTTTGTGGACTATACAAATGGCCTGGCATGAGATTTCTTCAATCCG 413
QY 268 GAGTAATCAAACTTTATTTTCTTAAAGACAGTATTTCTGATCTTCATGAT 327
D 414 GAGTAATCAAACTTTATTTTCTTAAAGACAGTATTTCTGATCTTCATGAT 473
QY 328 TCTGACTTTTCCATTTCAAAATCTTATGATGCAAACTGGGAACAGGACCTGAGAC 387
D 474 TCTGACTTTTCCATTTCAAAATCTTATGATGCAAACTGGGAACAGGACCTGAGAC 533
QY 388 TTTTGTGTCAAGTACCTCCGTCATATTTTATTTTCAATGTATACATTTTCATT 447
D 534 -TTTGTGTCAAGTACCTCCGTCATATTTTATTTTCAATGTATACATTTTCATT 592
QY 448 CTTGGGACTGATACATTCATGCTGACGAGAACACCAAGGCAATTTAAACATCAA 507
D 593 CTTGGGACTGATACATTCATGCTGACGAGAACACCAAGGCAATTTAAACATCAA 652
QY 508 CCCCCAAATCTCTGGGGGCTAAGATTTCTGTGTGCAATCTGGGATTCATGTTCTT 567
D 653 CCCCCAAATCTCTGGGGGCTAAGATTTCTGTGTGCAATCTGGGATTCATGTTCTT 712
QY 568 ACTCTCTTGGCTAATGATTTGACCAACAGCAGCCGAGAGACAAGATGTGAAG-A 626
D 713 ACTCTCTTGGCTAATGATTTGACCAACAGCAGCCGAGAGACAAGATGTGAAGAA 772
QY 627 AATCTCTTTTCTTAAATCAGAGTTGG--TCTAGTCTGGGATCAATAGTAATTCAT 684
D 773 AATCTCTTTTCTTAAATCAGAGTTGGGCTTAAAGTCTGGCCATTAATAGTAATTCAT 832
QY 685 CTGTCAGTCAATTTCTGATTAATTTCTTAATGTTATTTGATGTTATACATCATATAC 744
D 833 CTGTCAGTCAATTTCTGATTAATTTCTTAATGTTATTTGATGTTAT---CCCTCTTAC 889
QY 745 AAAAGACTGATACGCGTATACGTAAGAACGAGGGGTGTAGTAA 789
D 890 CAAAGAACTGGA-CGGTCATACGTAAGAACCAAGGGGTTGTGTA 933

```

```

RESULT 4
BG708321
LOCUS      680 bp      mRNA      linear      EST 07-MAY-2001
DEFINITION 602672256F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795236 5',
            mRNA sequence.
ACCESSION  BG708321
VERSION    BG708321.1  GI:13985546
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 680)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
            Toshiyuki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LHAM10677 row: 0 column: 13
            High quality sequence stop: 679.
            Location/Qualifiers

```

```

FEATURES
    source
        1..680
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NIH_MGC_96"
            /clone_1bp="NIH_MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /Note="Organ: brain; Vector: pBluescript (modified
            pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcag
            ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTT-3',
            size-selected for average insert size 2.3 kb and
            normalized to ROP 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, In preparation). Library
            constructed by M. Brownstein (NIH/NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      198 a      169 c      118 g      195 t
ORIGIN

```

```

Query Match      52.6%; Score 623.8; DB 12; Length 680;
Best Local Similarity 99.5%; Pred. No. 2e-125;
Matches 636; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

```

QY 28 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 87
D 33 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 92
QY 88 ATCAGGTACCAAGAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 147
D 93 ATCAGGTACCAAGAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 152
QY 148 CAGTCTGTGACACGAGACTACAAATCACCAGGTCCTCTCCOACTGCTACACTGT 207
D 153 CAGTCTGTGACACGAGACTACAAATCACCAGGTCCTCTCCOACTGCTACACTGT 212
QY 208 CCGTCTTTTGTGGACTATACAAATGGCCTGGCATGAGATTTCTTCAATCCG 267
D 213 CCGTCTTTTGTGGACTATACAAATGGCCTGGCATGAGATTTCTTCAATCCG 272
QY 268 GAGTAATCAAACTTTATTTTCTTAAAGACAGTCAATTTCTGATCTTCATGAT 327
D 273 GAGTAATCAAACTTTATTTTCTTAAAGACAGTCAATTTCTGATCTTCATGAT 332
QY 328 TCTGACTTTTCCATTTCAAAATCTTATGATGCCAAACTGGGAACAGGACCTGAGAC 387

```

```

|||||
Db 333 TGTGATTTTTCATTCATAAATTCCTTATGATGCAAACTGGGAAACGAGACACTGAGAAC 392
Oy 388 TTTTGTGTCAAGTACCTCCGTCATTTTATTTTACAAATGATATATGATTTTCATT 447
Db 393 TTTTGTGTCAAGTACCTCCGTCATTTTATTTTACAAATGATATATGATTTTCATT 452
Oy 448 CCTGGAGCTGATACATGATGCTGACAGAAAGACCAGGCCATTTAAACATCCAA 507
Db 453 CTGGGACATGATACATGATGCTGACAGAAAGACCAGGCCATTTAAACATCCAA 512
Oy 508 CCCCCAAATCTCTGGGGGGCTAAAGTCTCTCTGTCATCTGGGCAATCTGCTT 567
Db 513 CCCCCAAATCTCTGGGGGGCTAAAGTCTCTCTGTCATCTGGGCAATCTGCTT 572
Oy 568 ACTCTCTTGCCTAACATGATTTCTGACCAAGCAGCAGCAGACAAATGTGAA-CA 626
Db 573 ACTCTCTTGCCTAACATGATTTCTGACCAAGCAGCAGCAGCAGACAAATGTGAA-CA 632
Oy 627 AATGCTCTTTCCTTAAATCAGAGTTCGGTCTAGTCTGGC 665
Db 633 AATGCTCTTTCCTTAAATCAGAGTTCGGTCTAGTCTGGC 671

```

```

RESULT 5
LOCUS B1754688 1043 bp mRNA linear EST 25-SEP-2001
DEFINITION 603025383f1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5195900 5',
ACCESSION B1754688
VERSION B1754688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11490 row: e column: 21
High quality sequence stop: 733.
Location/Qualifiers
1..1043
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5195900"
/clone_11b="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains; age range 23-27 yo. Library is oligo-dr
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

```

BASE COUNT 340 a 228 c 205 g 265 t 1 others
ORIGIN
Query Match 52.0%; Score 616.8; DB 13; Length 1043;
Best Local Similarity 99.4%; Pred. No. 6.2e-124;
Matches 640; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

```

Oy 544 TGTGATCTGGGCAATTCATGTTCTTACTCTTTTGGCCTTAACATGATTTGACCAACAGCA 603
Db 13 TGTGATCTGGGCAATTCATGTTCTTACTCTTTTGGCCTTAACATGATTTGACCAACAGCA 70
Oy 604 GCCGAGACACAAAGATGTGAAGAAATGCTCTTCTTAAATCAGAGTTGGTCTACTG 663
Db 71 GCCGAGAGCAAGAAATGTGAAGAAATGCTCTTCTTAAATCAGAGTTGGTCTACTG 130
Oy 664 GCATGAATAGTAATTAATTAATGATGTCAGTCAATTTCTGATTAATTTCTTAATTTTAT 723
Db 131 GCATGAATAGTAATTAATTAATGATGTCAGTCAATTTCTGATTAATTTCTTAATTTTAT 190
Oy 724 TGTATGTTATACACTCATTACAAAGAACTGATACCGGTATAGTAAGACGAGGGTCT 783
Db 191 TGTATGTTATACACTCATTACAAAGAACTGATACCGGTATAGTAAGACGAGGGTCT 250
Oy 784 AGGTAAGTCCCGCAGGAAAAAGGTGACGTCAAAGTTTCATTATTCATTTGCTGATTTCT 843
Db 251 AGGTAAGTCCCGCAGGAAAAAGGTGACGTCAAAGTTTCATTATTCATTTGCTGATTTCT 310
Oy 844 TATTTGTTTGTCTCTTCCATTCTTCCGGAATTCCTACCCCTGAGCCAAACCGGGA 903
Db 311 TATTTGTTTGTCTCTTCCATTCTTCCGGAATTCCTACCCCTGAGCCAAACCGGGA 370
Oy 904 TGTCTTTGACTGACACTGCTGAAATTAATCTGTTCTATGTAAGAGACGACTGTGTT 963
Db 371 TGTCTTTGACTGACACTGCTGAAATTAATCTGTTCTATGTAAGAGACGACTGTGTT 430
Oy 964 AACTTCCTTAATAGCATGCGCTGATCCGTCATCTATTTTCTTCTTTCGCAAGTCTTCAG 1023
Db 431 AACTTCCTTAATAGCATGCGCTGATCCGTCATCTATTTTCTTTCGCAAGTCTTCAG 490
Oy 1024 AATTCCTGATTAATGATGCTGAGTGGCCCAATTCCTGCAATCTCTCTCCAGAGCA 1083
Db 491 AATTCCTGATTAATGATGCTGAGTGGCCCAATTCCTGCAATCTCTCTCCAGAGCA 550
Oy 1084 TAGGAAAAAAGAACAGAGATGTGTGACCCCAATGAAAGAGACTCCAAATTA 1143
Db 551 TAGGAAAAAAGAACAGAGATGTGTGACCCCAATGAAAGAGACTCCAAATTA 610
Oy 1144 ACTAAGGAATTAATTAATCTCTTCTGTCAGAACTCCTTAA 1187
Db 611 ACTAAGGAATTAATTAATCTCTTCTGTCAGAACTCCTTAA 654

```

```

RESULT 6
LOCUS BG716155 738 bp mRNA linear EST 08-MAY-2001
DEFINITION 602677640f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4800249 5',
ACCESSION BG716155
VERSION BG716155
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10690 row: p column: 10

```

High quality sequence start: 2
High quality sequence stop: 706.

FEATURES
Location/Qualifiers

1. 738

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1ib="IMAGE:4800249"

/clone_1ib="NIR_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI; XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to 10⁵. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carlnici, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIR_MGC Library."

BASE COUNT 210 a 184 c 126 g 217 t
ORIGIN

Query Match 50.5%; Score 600; DB 12; Length 738;

Best Local Similarity 95.3%; Pred. No. 2.9e-120;

Matches 672; Conservative 0; Mismatches 26; Indels 7; Gaps 5;

28 CTCGAGGCTGCAATTAACCTACTACTGATGATCAACCTCCAGATCAACAGTT 87
34 CTCGAGGCTGCAATTAACCTACTACTGATGATCAACCTCCAGATCAACAGTT 93
88 ATAGGTAACCAACAGAAATGCAAGCCGTCGACATCTCCTCGCCCTGGGAACAC 147
94 ATAGGTAACCAACAGAAATGCAAGCCGTCGACATCTCCTCGCCCTGGGAACAC 153
148 CAGTCTGTCACAGACATCAAAATCACCAGGCTCTCCACTGCTACACTGT 207
154 CAGTCTGTCACAGACATCAAAATCACCAGGCTCTCTCCACTGCTACACTGT 213
208 CCGTGTGTTGTTGACATTAACAAATGGCTGGCATGAGATTTCTTCAATCCG 267
214 CCGTGTGTTGTTGACATTAACAAATGGCTGGCATGAGATTTCTTCAATCCG 273
268 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTATTTGATCTTCATGAT 327
274 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTATTTGATCTTCATGAT 333
328 TCTGACTTTCATCAAAATCTTAGTATGCAAACTGGGAACAGACATGAGAAC 387
334 TCTGACTTTCATCAAAATCTTAGTATGCAAACTGGGAACAGACATGAGAAC 393
388 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTA--TATCAGTATTCAT 446
394 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTAATGATTTTCAT 453
447 TCTTGGGACGATTAATGATGCTGACAGAAAGACACAGGCCATTTAAACATCCA 506
454 TCTTGGGACGATTAATGATGCTGACAGAAAGACACAGGCCATTTAAACATCCA 513
507 ACCCCAA--AAATCTCTTGGGGGCTAAGATTTCTCTGTGTGATCTGGGATTCATTTTC 565
514 NACCCAAAGAAATCTCTTGGGGGCTAAGATTTCTCTGTGTGATCTGGGATTCATTTTC 573
566 TTAATCTCTTGGCTAATGATTTCTGACCAACAGGACGCGAGAGAACAA--TGTGA 623
574 TTAATCTCTTGGCTAATGATTTCTGACCAACAGGACGCGAGAGAACAAAGTGTGAC 633
624 AGAAATCTCTTGGCTAATGATTTCTGACCAACAGGACGCGAGAGAACAAAGTGTGAC 680
634 AGAAATCTCTTGGCTAATGATTTCTGACCAACAGGACGCGAGAGAACAAAGTGTGAC 693
681 ACAATCTGCAAGTATTTCTGATTAATTTCTTAATTTTATG 725
694 ACAATCTGCAAGTATTTCTGATTAATTTCTTAATTTTATG 738

RESULT 7

AO581896/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 609)

Zhao, S., Adams, M. D., Nierman, W., Malek, J., de Jong, P. and Venter

J. C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other GSSs: RPCI-11-433N7.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0208

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.bufrailo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.bufrailo.edu/ordering) or from

Research Genet. cs (info@resgen.com). BAC end search page:

http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. 609

/organism="Homo sapiens"

/db_xref="GDB:7665206"

/db_xref="taxon:9606"

/clone="RPCI-11-433N7"

/clone_1ib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCI11 Human Male BAC Library"

BASE COUNT 204 a 104 c 122 g 179 t

ORIGIN

Query Match 50.5%; Score 599.6; DB 17; Length 609;

Best Local Similarity 99.3%; Pred. No. 3.6e-120;

Matches 602; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

261 AAATCCGAGTAATCAACTTATTTATTTCTTAAGAACACAGTATTTGATCTTC 320
609 AAATCCGAGTAATCAACTTATTTATTTCTTAAGAACACAGTATTTGATCTTC 550
321 TCATGATTCGACTTTCATCAAAATCTTAGTATGCAAACTGGGAACAGGACAC 380
549 TCATGATTCGACTTTCATCAAAATCTTAGTATGCAAACTGGGAACAGGACAC 490
381 TGAGAACTTTGTGTGCAAGTACCTCCGTCATATTTTATTCACATGTAATTCAGTA 440
489 TGAGAACTTTGTGTGCAAGTACCTCCGTCATATTTTATTCACATGTAATTCAGTA 430
441 TTTTATTCCTGGGACGATTAATGATGCTGACAGAAAGACACAGGCCATTTAAA 500
429 TTTTATTCCTGGGACGATTAATGATGCTGACAGAAAGACACAGGCCATTTAAA 370
501 CATCAACCCCAAAATCTCTTGGGGCTAAGATTTCTCTGTGTGATCTGGGACATCA 560
369 CATCAACCCCAAAATCTCTTGGGGCTAAGATTTCTCTGTGTGATCTGGGACATCA 310

QY 561 TGTCTTACTCTCTTGGCCCTACATGATTCGACCAACAGGAGCCGAGACAGCAAGATG 620
 DB 309 TGTCTTACTCTCTTGGCCCTACATGATTCGACCAACAGGAGCCGAGACAGCAAGATG 250
 QY 621 TGAAGAAAGCTCTTCCCTTAATGAGAGTGGGCTACTGCGCATGAAATAGTAAAT 680
 DB 249 TGAAGAAAGCTCTTCCCTTAATGAGAGTGGGCTACTGCGCATGAAATAGTAAAT 190
 QY 681 ACATCTGTCAGTCATTTCTGATTAATTTCTTAATTTGTAATGTAATGTAATGTAAT 740
 DB 189 ACATCTGTCAGTCATTTCTGATTAATTTCTTAATTTGTAATGTAATGTAATGTAAT 130
 QY 741 TTACAAAGAACTGTACGCGTCATACGTAAGAACGAGGGGTAGTAAAGTCCACAGA 800
 DB 129 TTACAAAGAACTGTACGCGTCATACGTAAGAACGAGGGGTAGTAAAGTCCACAGA 70
 QY 801 AAAGGTGAGGTCAAGTTTCATTAATGATGCTGATCTTATTTTGTGTTCTCT 860
 DB 69 AAAGGTGAGGTCAAGTTTCATTAATGATGCTGATCTTATTTGTTGTTCTCT 10
 QY 861 TCCATT 866
 DB 9 ACCATT 4
 RESULT 8
 LOCUS B0893760 971 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOURT_8482519_lupski_dorsal_root_ganglion Homo sapiens CDNA
 accession B0893760
 version B0893760.1 GI:22285774
 keywords EST.
 source human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 971)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M13577 row: m column: 06
 High quality sequence stop: 607.
 Location/Qualifiers
 1..971
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6186029"
 /clone_lib="lupski_dorsal_root_ganglion"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 directionally cloned using the following adaptors:
 5'-TCGACCCACGCTCCG-3' and
 5'-GACTAGTTCTGATCGAGCGCGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 299 a 236 c 172 g 264 t
 ORIGIN
 Query Match 48.6%; Score 577.2; DB 14; Length 971;
 Best Local Similarity 97.7%; Pred No. 2.5e-115;
 Matches 596; Conservative 0; Mismatches 13; Indels 1; Gaps 1;
 QY 28 CTGCAAGCTGCAATAACTACTACTGATGATGATCAATTCMAACCTCCAGAAATCAACAGTT 87
 DB 148 CTGCAAGCTGCAATAACTACTACTGATGATGATCAATTCMAACCTCCAGAAATCAACAGTT 207
 QY 88 ATCAGGTAAACCAAGAAATGCAAGCCGTCGACAAATCTCACCCTGCGCTGGGAACAC 147
 DB 208 ATCAGGTAAACCAAGAAATGCAAGCCGTCGACAAATCTCACCCTGCGCTGGGAACAC 267
 QY 148 CAGCTGTGTCACAGAGACTACAAATACCCAGGTCCTCCACAGTCTCTACAGT 207
 DB 268 CAGCTGTGTCACAGAGACTACAAATACCCAGGTCCTCCACAGTCTCTACAGT 327
 QY 208 CCTGTTTTTTGTTGAGCTTATGCAAAATGGCCTGGCGATGAGGATTTCTTCAAAATCG 267
 DB 328 CCTGTTTTTTGTTGAGCTTATGCAAAATGGCCTGGCGATGAGGATTTCTTCAAAATCG 387
 QY 268 GAGTAATCAAACTTATTTATTTCTTAAGACACAGTCATTTCTGATCTTCTATGAT 327
 DB 388 GAGTAATCAAACTTATTTATTTCTTAAGACACAGTCATTTCTGATCTTCTATGAT 447
 QY 328 TCTGACTTTTCCATTCAAAATCTTATGATGCGCAAACTGGGAGACAGCAGCTGGAAC 387
 DB 448 TCTGACTTTTCCATTCAAAATCTTATGATGCGCAAACTGGGAGACAGCAGCTGGAAC 507
 QY 388 TTTTGTGTGTCAGTACCTCGCTCATATTTATTTACAAATGTAATACGATTTTCAT 447
 DB 508 TTTTGTGTGTCAGTACCTCGCTCATATTTATTTACAAATGTAATACGATTTTCAT 567
 QY 448 CCTGGAGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
 DB 568 CCTGGAGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
 QY 508 CCCCAAAAATCTTGGGGGCTAAGATTCCTCTGTCATCTGGGCAATCATGTTCT 567
 DB 628 CCCCAAAAATCTTGGGGGCTAAGATTCCTCTGTCATCTGGGCAATCATGTTCT 687
 QY 568 ACTCTCTTGGCTTAACATGATTTCTGACACACA-GGCAGCCGAGAGCAAGATGGAAGA 626
 DB 688 ACTCTCTTGGCTTAACATGATTTCTGACACACA-GGCAGCCGAGAGCAAGATGGAAGA 747
 QY 627 AATGCTCTT 636
 DB 748 AATGCTCTT 757
 RESULT 9
 LOCUS BM987216 511 bp mRNA linear EST 25-MAR-2002
 DEFINITION UT-H-CO0-aq1-c-12-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens CDNA clone
 accession BM987216
 version BM987216.1 GI:19706605.
 keywords EST.
 source human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 511)
 AUTHORS NCI-CGAP http://www.nci.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
SOLTA+Yes.

FEATURES	Location/Qualifiers
source	1. .511

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE: 3104254"
/clone_id="NCI_CGAP_Sub9"
/tissue_type="mixed"
/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
Cholonic mucosa with Crohns disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, cervical
adenosquamous carcinoma, ligament cells, prostate
carcinoma, Bladder carcinoma, brain oligodendroga
; NCI_CGAP_Sub9 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are CCGC, ATGC, GGCGC,
GGAAG, TAGC, TTAGC, ATGG, AGACA, ATCAC. For additional
information, contact: Bento Soares, bento-soares@iowa.edu
TAG_LIB=UI-H-COO
TAG_TISSUE=Brain Oligodendroga
TAG_SEQ=ATTCAC"

```

	Query Match	Similarity	92.8%	Score 507.8	DB 14	Length 511
	Best Local	Similarity	99.6%	Pred. No. 3.3e-100		
	Matches	509	Conservative	0	Mismatches	2
					Indels	0
					Gaps	0
QY	281	TTTATTTATTTTCTTTAGAGACACAGCTCATTTTGTGATCTTCATGATATTTGTACTTTTCCA	340			
Db	1	TTTTTTTTTTTTTCTTAAGAACAGACTCATTTTGTGATCTTCATGATATTTGTACTTTTCCA	60			
QY	341	TTCAAAATTTCTTAGTGATGCCAAACTGGGAGACAGACACATGAGAACTTTTGTGTCAA	400			
Db	61	TTCAAAATTTCTTAGGAGGCCAAACTGGGAGACAGACACATGAGAACTTTTGTGTCAA	120			
QY	401	GTTACCTCCGTCATTTTATTTTACCAATGTTATACAGATTTTCAATTCCTGGGACCTGANA	460			
Db	121	GTTACCTCCGTCATTTTATTTTACCAATGTTATACAGATTTTCAATTCCTGGGACCTGANA	180			
QY	461	ACTATCGATCGCTACAGAGACACACAGGCCATTTTAAACATCCAACCCAAAAATCTC	520			
Db	181	ACTATCGATCGCTACAGAGACACACAGGCCATTTTAAACATCCAACCCAAAAATCTC	240			
QY	521	TTGGGGGCTAGATTTCTCTCTGTTGTCATCTGGCATTCATATTTCTTACTCTCTTGGCT	580			
Db	241	TTGGGGGCTAGATTTCTCTCTGTTGTCATCTGGCATTCATATTTCTTACTCTCTTGGCT	300			
QY	581	AACATGATTTGACCAACAGCGGACGCGAGAGACAAGAAATGGAAGAAATGCTTTCCTT	640			
Db	301	AACATGATTTGACCAACAGCGGACGCGAGAGACAAGAAATGGAAGAAATGCTTTCCTT	360			
QY	641	AAATCAGAGTTCGCTCTAGCTGCGCATGCAAAATAGTAATTTACATCTGTCAAGTCAATTTTC	700			
Db	361	AAATCAGAGTTCGCTCTAGCTGCGCATGCAAAATAGTAATTTACATCTGTCAAGTCAATTTTC	420			
QY	701	TGGATTAATTTCTTAATTTGTATTTGTATGTTATACACTCATTTACAAAAGACTGTCGGG	760			

Db	421	TGATTAATTTCTTAATTTGTTATTTGTATGTATTATCACTCATTTACAAAGACTGTACGG	480
QY	761	TCATACGTTAAGACGAGGGGCTGTAGCTAAAG	791
Db	481	TCATACGTTAAGACGAGGGGCTGTAGCTAAAG	511

RESULT	10
AV718096	
LOCUS	706 bp mRNA linear EST 16-OCT-2000
DEFINITION	AV718096 FHTA Homo sapiens cDNA clone FHTABADZ 5', mRNA sequence.
ACCESSION	AV718096
VERSION	AV718096.1 GI:10815248
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1 (basses 1 to 706) Xiao, H., Peng, Y., Song, H., Gu, Y., Yang, Y., Gao, G., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Chen, Z. and Han, Z.	Homo sapiens CDNA FHTA clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@hgc.sh.cn

FEATURES	Location/Qualifiers
source	1. .706

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FHTABA02"
/clone_id="FHTA"
/tissue_type="hypothalamus"
/dev_stage="Petal"
/lab_host="BM25.8"
/note="Vector: pRR16x2; Site:1: sf11a; Site:2: sf11b1"

```

Query Match	42.6%	Score 505.4	DB 10	Length 706
Best Local Similarity	86.6%	Pred. No. 11e-99		
Matches 602	Conservative	0	Mismatches 38	Indels 55
				Gaps 2
Qy	36	TGCATTAACCTACTACTGTAATGATGATCAATTCGAACCCCTGCAGATCAACAGTTATACGGTA	95	
Db	64	TGCATTAACCTACTACTGATGATGATCAATTCGAACCCCTGCAGATCAACAGTTATACGGTA	123	
Oy	96	ACCAACAAGAAATGCAAGCCGTGCACAATCTGACCTCTGGCTGGGAACACA -GTCTG	154	
Db	124	ACCAACAAGAAATGCAAGCCGACACAACAACCTGACCTCTGGCTGTTAACACAGATGCTG	183	
Oy	155	TGCACACAGAGACTACAAAATTCACCCAGGTCCTCTTCCACTGCTCTACACTGCTCTT	214	
Db	184	TGCACACAGAGACTACAAAATTCACCCAGGTCCTCTTCCACTGCTCTCTACACTGCTCTT	243	
Oy	215	TTTGTGGACTATACAAAATGGCCGCGGATGAGGATTTCTTTCAAATCCGGACTAA	274	
Db	244	TTTGTGGACTATACAAAATGGCCGCGGATGAGGATTTGTTCAAAATCCGGACTAA	303	
Oy	275	TCAAATTTATTTATTTCTTAAAGACACAGTCATTTGATCTTCTCATGATTCGACT	334	
Db	304	TCAAATTTATTTATTTCTTAAAGACACAGTCATTTGATCTTATCATGATCTGACT	363	
Oy	335	TTTTCAATTCAAAATTTCTTAAGTATGCCAAATGGGACAAGCAACCTGAGAACTTTTGTG	394	
Db	364	TTTTCAATTCAAAATTTCTTAAGTATGCCAAATGGGACAAGCAACCTGAGAACTTTTGTG	423	

QY	395	TGTCAGATTACCTCCGTCATATTTATTTACAAATGATATACGATATTTCAATCCGGGA	454
Db	424	TGTCAGTTACTCGGTATATTTATTTCAAAATGATATACGATATTTGATTCCTGGGA	483
QY	455	CTGATTAACATATCGATCCGCTACCGAAGACACACAGGCCATTTTAAACATCCAAACCCAAA	514
Db	484	CTGATTAACATATCGACCGGTACCGAAGACACACAGGCCATTTTAAAGCAATCCAAACCCCA --	541
QY	515	AATCTCTTGGGGGGCTAAGATTCTCTCTGTGTGCATCTGGGCAATCATGTCTTACTCTCT	574
Db	542	-----TACTCTGT	549
QY	575	TTTGCTAATCATGATTTCTACCAACAAGGACGCCGGAAGACACAAATGTCAAGAAATGCTCTT	634
Db	550	ATGCTTAATCATGACCTTACCAACAAGGACGCCGGAAGACACAAATGTGTGAAGTAAATGCTTA	609
QY	635	TTTCCTTAAATAGAGTTGGTCTACTCTGCGCATGGAATAGTAAATTTACATCTGTCAGATG	694
Db	610	TTTCCTTAAATCAGACCTCGCTCTACTCTGCGCATGGAATAGTAAADATCTGNCNAAGTC	669
QY	695	ATTTTCTGATTAATTTTCTAATTTTATTTGATG	729
Db	670	ATTTTCTGATTAATTTGCTTAATTCCTATG	704

RESULT 11	
LOCUS	BB624828
DEFINITION	BB624828 674 bp mRNA linear EST 31-AUG-2001 BB624828 RIKEN full-length enriched, adult male colon Mus musculus cDNA clone 9030207C05 5', mRNA sequence.

INCIDENT 220478Z
VERSION BB624828.1 GI:15398391
KEYWORDS EST.

SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
1 (bases 1 to 674)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogomphidae; Muridae; Mus-
1 (bases 1 to 674)

TITLE
 JOURNAL
 COMMENT
 Contact: yoshihide Hayashizaki
 Unpublished (2001)

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-1-72 Suheiho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216

URL: genome-resescrc.riken.go.jp,
<http://genome.gsc.riken.go.jp/>
 Carinacci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wael, K., Fujikawa, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E.,
 Wataniki, M., Itoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
 S., Kawai, I., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.

10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a

FEATURES
source
nonredundant cdna library
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arkawa,T.,
Isii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, 172-186 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.
Location/Qualifiers
1..674

```

BASE COUNT
ORIGIN
176 a 171 c 122 g 205 t
BamHI"
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9030207C05"
/clone_lib="RIKEN full-length enriched, adult male colon"
/sex="male"
/tissue_type="colon"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGAGAGAGAGATCCACAGACCTCTTTTCTTTTCTTTTCTVN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round.of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGATCTCGAGTTAATTAATATCCCCCCCCCC 3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Salt; 3' end: BamHI"

```

Query Match	37.9%	Score 449.4	DB 10	Length 674
Best Local Similarity	84.1%	Pred. No. 1,66-87		
Matches 507	Conservative	0	Mismatches 96	Indels 0
			Gaps	
QY	114	CGGTGCAATCTCACTCTGGCGCTGGGAACACACAGTCTGTGCACGAGAGCTACAAAA	173	
Db	72	CCTGAGCAAAATCACCCCTTCCTCCCTGGGACACACACCTGTGGTGAAGACATCAAGA	131	
QY	174	TCACCCAGAGTCCCTTCCCACTGCCTCAACACTGTCTCTGTCTTTGGTAGATTATCAAA	233	
Db	132	TCACCCAGGTTCTCTTCCCATGTGCTGTAACACCGTCTGTCTTTGGCGGCTATCAAGA	191	
QY	234	ATGGCCCTGGCGCATAGAGATTTCTTTCAATCCGAGTAATCAACTTTATATATTTTC	293	
Db	192	ACACCTTTGGCAATGAGATTTTCTTTTCAGATCCGACAGTAATCAACCTTCATCATTTTC	251	
QY	294	TTAAGAACACAGTATTTCTGATCTCTGCATGATTCGTACTTCTTCATTCAAATTCCTA	353	
Db	252	TTAAGAACAGGTCATCTGTGATCTGATCTAGTAATGATTCTTAACCTTTCCATTTAAATTCCTA	311	
QY	354	GTGATGCCAAACTGTGGAAACAGGACCACTGAGAACTTTTGTGTGCAAAGTTACCTCCGTCA	413	
Db	312	GTGATGTGTAACCTGTGGAGCGGGGCTCTGGAACCTTGGTGTGCCAAGTTACTTCAGTCA	371	
QY	414	TATTTTATTTACAAATATATATCAATATTTTCATTTCTGTGGACATGATACTCATGATGCT	473	
Db	372	CATTTTATTTTACAAATGATATATCATGATATATGCTTCCGTGGGGTGTATATCAACTATACCGCT	431	
QY	474	ACCGAAGACCAACGAGGCATTTTAAACATCCAAACCCCAAAATCTCTTGGGGGCTAAGA	533	
Db	432	ACCGAAGACCAACGAGGCATTTTAAACATCCAAACCCCAAAATCTCTTGGGGGCTAAGA	491	

QY	534	TTCTCTCTGTTGTAATCTCGGCATCTCAGTTCTTACTCTCTTGGCTTAACATGATTTGCA	593
Db	492	TTCTTTCTGTTGTACTCTCGGCCTTCACGTTCTTATTTCACTGCTTACATGATTTGCA	551
QY	594	CCAACAGGACCCGAGACACAAGAAATGTGAAGAAATGCTCTTTCCTTAAATCAGATTGCG	653
Db	552	CCAACAGGACCCCAAAAGATTAAGGACGTAACAAAGTTCTTCTTAAGTCAGAGTTTG	611
QY	654	GTCAGTGTGCGAATTAATTAATTCATCTGCATGCAAGTCAATTTTCGATTAATTTCT	713
Db	612	GTCAGTGTGCGAATTAATTAATTCATCTGTGCAAGTCAATTTTCGATTAATTTT	671
QY	714	TAA	716
Db	672	TAA	674
RESULT 12			
LOCUS	BM926255	1681 bp	mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT 6600668 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764646		
ACCESSION	BM926255		
VERSION	BM926255.1	GI:19376634	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Ekaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	Nih-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabos-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LHAM12818 row: g column: 15 High quality sequence stop: 512.		
FEATURES			
source	Location/Qualifiers		
	1..1681		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5764646"		
	/clone_1lb="NIH_MGC_114"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-Sport6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH-MGC Library."		
BASE COUNT	513 a	309 g	407 t
ORIGIN	13 c		13 others
Query Match	37.3%	Score 442.4	DB 14; Length 1681;
Best Local Similarity	95.28;	Pred. M4.7e-86;	
Matches 501; Conservative	0;	Mismatches 16;	Indels 9; Gaps 4;
QY	28	CTGCAGGCTGCAAACTACTACTACTGATTCATCAACCCCTCCAGAAATCAACAGTT	87
Db	138	CTGCAGGCTGCAAACTACTACTACTGATTCATCAACCCCTCCAGAAATCAACAGTT	197
QY	88	ATCAGTACCAACAAGAAATGCAAGCCGTGACATCTCACCTTGCGCCTGGAGAC	147

Db	198	ATCA - GTACACACAGAAATGCAAGCCGTGCACAACTCTGCGCCCTGGGAACAC	256
QY	148	CAGTGTGTGCAACGAGACATACAAATACACCAGGTCCTCTCCAGCTGTACACTGT	207
Db	257	CAGTGTGTGCAACGAGACTACAAATACACCAGGTCCTCTCCAGCTGTACACTGT	316
QY	208	CCGTGTTTTGTGGACATTCACAAATGGCCCTGGCGATGAGATTTTCTTCAATCCG	267
Db	317	CCTGTTTTTTGTGGACTTATACAAATGGCCCTGGCGATGAGATTTTCTTCAATCCG	376
QY	268	GAGTAATCAAACTTTATTTTCTTAAAGAACAGTCATTTTCGATCTTCATGAT	327
Db	377	GAGTAATCAAACTTTATTTTCTTAAAGAACAGTCATTTTCGATCTTCATGAT	436
QY	328	TCTGACTTTTCATTCAAAAATTTCTTAGTGATGCCAAACTGGGACACAGCACTGAGAC	387
Db	437	TCTGACTTTTCATTCAAAAATTTCTTAGTGATGCCAAACTGGGACACAGCACTGAGAC	496
QY	388	TTTTGTGTGCAAGTATACCTCGCTCATTTTTTATTTCAAAATGTATATCAGATTTTCAT	447
Db	497	TTTTGTGTGCAAGTATACCTCGCTCATTTTTTATTTCAAAATGTATATCAGATTTTCAT	556
QY	448	CTGTGGACTGTATACATATCGATGCGTACAGAACACACACAG--CCATTTAAACATC	504
Db	557	CTGTGGACTGTATACATATCGATGCGTACAGAACACACACAGCCATTTTAAACATC	616
QY	505	CAACCCCAAAAA-----TCTCTGGGGGCTAAGATTCCTCTGTTG	545
Db	617	CAACCCCAAAAACTCTCTTGGGGGCTCAAAATTTCTCTTCTG	662
RESULT 13			
LOCUS	BG919208	877 bp	EST 05-JUN-2001
DEFINITION	60281803F1 NCL_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4947057		
ACCESSION	BG919208		
VERSION	BG919208.1		
KEYWORDS	GI:14299684		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 877)		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.nih.gov Plate: LHAM10896 row: e column: 10 High quality sequence stop: 815. location/Qualifiers 1. 877 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone_image="IMAGE:4947057" /clone_id="NCL_CGAP_Mam6" /sex="female, virgin" /tissue_type="infiltrating ductal carcinoma" /dev_stage="5 months" /lab_host="DH10B" /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"		

ORIGIN	BASE COUNT	233 a	224 c	169 g	251 t
Query Match	35.7%;				
Best Local Similarity	82.2%;				
Matches 512;	Conservative 0;				
	Mismatches 108;				
	Indels 3;				
	Gaps 2;				
OY	114	CCGTGACAAATCTCCCTCTGCGCTGGGGAACACAGTCTGTGCACACGAGCTACAAA	173		
Db	242	CCTGACCAATATACACCTCTCTCCCTGGGACAGACCTCTGCTGTACAGACTACAAA	301		
OY	174	TCACCCAGGCTCTCTCCACCTCTACACGTCTGCTGTCTGTGTGTGGACTATACAA	233		
Db	302	TCACCCAGGCTCTCTCCACCTCTGCTGTACACGCTCTCTCTCTCTGCTGTACACGA	361		
OY	234	ATGCGCTGGGCAATGAGATTTTCTTTCAATCCGAGTAATCAATCAATTAATTTTC	293		
Db	362	ACAGCTTGCAATGAGATTTTCTTTCAATCCGAGTAATCAATCAATTAATTTTC	421		
OY	234	TTAAGAACACAGTCTATTTCTGATCTTCTCATCTATTTCTGATTTTCCATTTCAATTTCTTA	353		
Db	422	TTAAGAACACAGTCTATTTCTGATCTTCTCATCTATTTCTGATTTTCCATTTCAATTTCTTA	481		
OY	354	GTGATGCCAAATCGGGAACAGGACACAGTGAAGTCTGTGTGAAGTACTCCGTCA	413		
Db	482	GTGATGCCAAATCGGGAACAGGACACAGTGAAGTCTGTGTGAAGTACTCCGTCA	541		
OY	414	TATTTATTTTCAATGATATATACATATTTTCAATTTCTGCTGGACTGAATATGACATCGT	473		
Db	542	CATTTATTTTCAATGATATATATACATATATGCTTCTGCTGGGATGATTAACATATGACGCT	601		
OY	474	ACCAGAACACACACAGGCTATTTAAACATCCACCCCAAAATCTCTTGGGGCTAGA	533		
Db	602	ACCTGAACACACACAGGCTATTTAAACATCCACCCCAAAATCTCTTGGGGCTAGA	661		
OY	534	TTCTCTGTTGTCAATCTGGGACATCATAGTCTTCTGCTGCTTGTGCTTAACATGATCTGA	593		
Db	662	TTCTCTGTTGTCAATCTGGGACATCATAGTCTTCTGCTGCTTGTGCTTAACATGATCTGA	721		
OY	594	CCAACAGGACCCGAGAGACACAAGATGTGAAGAAA - TGCTTTTCCCTTAATACAGAGTT	651		
Db	722	CCAACAGGACCCGAGAGACACAAGATGTGAAGAAA - TGCTTTTCCCTTAATACAGAGTT	781		
OY	652	CGGTCTAGTCTGCGCATGAAATAGTAAATTTACATCTGTCAAGTCAATTTCTGGATTAAATT	711		
Db	782	TGCTCTAGTCTGCGCATGAAATAGTAAATTTACATCTGTCAAGTCAATTTCTGGATTAAATT	840		
OY	712	CTTAATGTTATTTGATGATTATA	734		
Db	841	TTTAATGTTATTTGATGATTATA	863		
RESULT 14					
LOCUS	AA274112	542 bp	mRNA	linear	EST 28-MAR-1997
DEFINITION	vb92d04.t1 Soares mouse 3BDM5 Mus musculus CDNA clone IMAGE:764455				
	5' similar to TR:G285995 G285995 ORF, COMPLETE CDS. ;, mRNA				
sequence.					
ACCESSION	AA274112				
VERSION	AA274112.1	GI:1912557			
KEYWORDS	EST.				
ORGANISM	house mouse.				
SOURCE	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 542)				
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,				
	Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,				
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,				
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and				
	Waterston,R.				
TITLE	The WashU-HMI Mouse EST Project				
JOURNAL	Unpublished (1996)				

COMMENT	Contact: Maria M/Mouse EST-Project WashU-HMHI Mouse EST Project Washington University School of MedicineP 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LNL : contact the IMAGE Consortium (info@image.lnl.gov) for further information. MI:465375 Possible reversed clone; similarity on wrong strand Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 504.
FEATURES	Location/Qualifiers
SOURCE	1. 542 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:764455" /clone_lib="Soares mouse 3MBDS" /sex="male" /tissue_type="Spleen" /dev_stage="4 weeks" /lib_host="DH10B" /note="Vector: pT7r3D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCATCTGAAGTCGAGCGGCCGGCCGTATTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7r3 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	139 a 125 c 103 g 175 t
ORIGIN	
Query Match	34.3%; Score 407.6; DB 9; Length 542;
Best Local Similarity	84.5%; Pred. No. 2e-78;
Matches 456:	Conservative 0; Mismatches 84; Indels 0; Gaps 0;
OY 375	GACCATTGAGAAGACTTTGTGTGCAGTAAGTACCCTCCGTCANATTTATTTACAAATGATA 434
Db 1	GGCCTCGAGAACCTTGGGTGCAAGTAGTTACTGCACATGTTATTTACAATGTATA 60
OY 435	TCACTAATTTCTTCTCTGGGACATGATACGTACGATCCGTACCAGAAAGCACCGCCAT 494
Db 61	TCAGTATATCGTCTCTGGGGTGAATMACCATTTGACCGGTACTGAAAAGCACCGCCAT 120
OY 495	TTAAACATCCAAACCCCACAAAAATCTCTGGGGGGTAAAGATTCTCTGTGTGCATCTGGG 554
Db 121	TTAAACGTCGACGCCACGACATCTCTTGGGTGCAGAAATCTTCTGTGTGCATCTGGG 180
OY 555	CATTCAATGTTCTACTCTCTTGGCTTAACATGATTTTGACACAGCAGCGGAGACA 614
Db 181	CCITTCATGTTCTTATTTCACTGCTTAACATGATTTTCACACAGGAGGCCAAAGATA 240
OY 615	AGAAATGGAGAAANGCTTCTTCTTAANCAGAGTCCGTCCTGTCTGCGATCAAANAAG 674
Db 241	AGGACGTAACAAAATGTCTTTCTTTAAAGTAGAGTTGGCTGTAGTTGGACACAAAATAG 300
OY 675	TAAATTCATCTGTCAAGTCATTTTCTTGAGATTAATTTCTTAAATGTATATGTATATATA 734
Db 301	TCATATTCATCTGTCCAGTCATTTTCTTGAGATTAATTTTAAATGTATATCTTTGTATATA 360
OY 735	CACATCATTTACAAAAGAACTGTACCGGTCAATACGTAAAGACGAGSGGTAGTAAGTCC 794
Db 361	GCCCTCATTCACAAAAGAACTGTACGGTCTTATGTACAGAACAGGSGGTCAACCAAGTTC 420
OY 795	CCAGGAAAAAGGTCAGAGTCAAAGTTTTCATTAATCATTTGCGTATNTCTTTATTTGGTTTG 854
Db 421	CCAAGAGAAAGTAAAGTCACAGGTTTTCATCATCATATGCTGTATTTCTTTATTTGGCTTTG 480

QY	855	TTCCCTTCATTTTCCCGGAATTCCTTACACCCCTAGCAAAACCCGGATCTTGTACT	914
Db	481	TTCCCTTCACCTTTCACAGGATTCCTTCACCCCTAGCAGACTCGGGCCCTTTGACT	540
QY	915	GC 916	
Db	541	GC 542	
RESULT 15			
LOCUS	Bf677440		
DEFINITION	602087191.F1 NIH_MGC_83 Homo sapiens CDNA clone IMAGE:4251263 5',		
ACCESSION	Bf677440		
VERSION	Bf677440.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: rs9ab@rsfmail.nih.gov		
	Tissue Procurement: CLONTECH Laboratories, Inc.		
	CDNA Library Preparation: CLONTECH Laboratories, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LNCMI074 row: e column: 24		
	High quality sequence stop: 541.		
FEATURES			
source	Location/Qualifiers		
	1..915		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4251263"		
	/clone_lib="NIH_MGC_83"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/note="Organ: prostate; Vector: pNNR-LIB (Clontech);		
	Site_1: SfilI (ggcgccctggcc); Site_2: SfilI (ggcattatggcc		
	; 5' and 3' adaptors were used in cloning as follows: 5'		
	adaptor sequence: 5'-CACGGCAATTAAAGGC-3' and 3' adaptor		
	sequence: 5'-ATTCTAGAGCCGACGGCCGACATC-dt(30)BN-3',		
	(where B = A, C, or G and N = A, C, G, or T). Average		
	insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies		
	contained inserts by PCR. This library was enriched for		
	full-length clones and was constructed by Clontech		
	Laboratories (Palo Alto, CA)."		
BASE COUNT	281 a 231 c 190 g 213 t		
ORIGIN			
Query Match	30.9%; Score 366.8; DB 12; Length 915;		
Best Local Similarity	91.0%; Pred. No. 1.4e-69;		
Matches 457; Conservative	0; Mismatches 37; Indels 8; Gaps 6;		
QY	28	CTGCAAGCGCTGCATTAACACTACTGATGATATCAATCAACCCCTGCAGATCAACATTT	87
Db	166	CTGCAAGCGCTGCATTAACACTACTGATGATATCAATCAACCCCTGCAGATCAACATTT	224
QY	88	ATCAGGTAACCAAGAAATGACAGCGGTGCACATCTACCTGGCGCTGGGAACG	147
Db	225	ATCAGGTAACCAAGAAATGACAGCGGTGCACATCTACCTGGCGCTGGGAACG	284
QY	148	CAGTCTGTGACACAGAGACTACAAATGACACCGAGTCTCTTCCCACTGCTTACACTGT	207
Db	285	CAGTCTGTGACACAGAGACTACAAATGACACCGAGTCTCTTCCCACTGCTTACACTGT	344
QY	208	CCTGTTTTTTTGTGGACTTATCATCAAAATGCGCTGGCGATGAGGATTTCTTCAAAATCCG	267

Db	345	CGTG-TTTTTGTGGGAGCTTATCACAAATGGCC-T-GCGATGGAGATCTTTTCAAATCCG	4020
Oy	268	GAGTAAATCAAACTTTATTTATTTTCTTAAAGACACAGTCATTTCTGATCTTCTCATGAT	327
Db	403	GAGTAAATCAAACTTTATTTATTTTCTTAAAGAACACAGTCATTTCTGATCTTCTCATGAT	4620
Oy	328	TCTGCACTTTTCATTCAAAATCTTTAGAGATGCCAAACTGGGAACAGAGACACGAGAAC	387
Db	463	TCTGCACTTTTCATTCAAAATCTTTAGAGATGCCAAACTGGGAACAGAGACACCTGAGAAC	5222
Oy	388	TTTTTGTGTCAAGTAACTCCCTCATATTTTATTTTCAACATGTATTTACAGTATTTTC	4455
Db	523	TTTTTGTGTCAAGTAACTCCCTCATATTTTATTTTCAACAGTATTTATTCAGTATTTTCAT	5822
Oy	446	TTTCTGGGACTGA-TAACTATCGATCGGTACACAGAACACACAGGCCATTTTAAACATC	5040
Db	583	TCTCTGGGACACTGATCACTATTCGTTCCGTACAGAACACAC - AGSCATTTTAAACATC	6400
Oy	505	CAACCCCAAAAATCTCTTGGGG 526	
Db	641	CAAGCCCAAAAATTTTGTGGGG 662	

Search completed: December 11, 2002, 07:39:36
Job time : 2208 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 22:55:05 ; Search time 3226 Seconds

(Without alignments)
10708.317 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187

Sequence: 1 ggtaccggggagatcctctag.....tgtgttcagacgtcttaaa 1187

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1160	97.7	1424	9 BC017898	BC017898 Homo sapi
2	1153.4	97.2	1332	9 AF313449	AF313449 Homo sapi
3	1152	97.1	1152	9 AF321815	AF321815 Homo sapi
4	1110.4	93.5	1951	9 AB056385	AB056385 Macaca fa
5	1108.8	93.4	2266	9 AB062981	AB062981 Macaca fa
6	1097	92.4	328187	2 AC117393	AC117393 Homo sapi
7	1093.8	92.1	1143	9 AF311401	AF311401 Sequence
8	1093.8	92.1	1143	9 AF310685	AF310685 Homo sapi
9	1093.8	92.1	173270	9 AC024886	AC024886 Homo sapi
10	1027.4	86.6	1029	6 AX369349	AX369349 Sequence
11	1027.4	86.6	1029	9 AB052684	AB052684 Homo sapi
12	1025.8	86.4	1029	9 AB083596	AB083596 Homo sapi
13	732.2	61.7	2151	10 BC027381	BC027381 Mus muscu
14	732.2	61.7	2422	10 BC025428	BC025428 Mus muscu
15	732.2	61.7	211964	2 AC122038	AC122038 Mus muscu
16	729.8	61.5	1044	6 AX369354	AX369354 Sequence
17	729.8	61.5	1047	10 MM0312130	AF313450 Mus muscu
18	701.4	59.1	1254	10 AF313450	AF313450 Rattus no
19	695	58.6	179889	2 AC096965	AC096965 Rattus no
20	466	39.3	466	6 AX369353	AX369353 Sequence
21	407.6	34.3	542	6 AX413092	AX413092 Sequence
22	307.2	25.9	1002	6 E49123	E49123 Novel G pro
23	307.2	25.9	1002	6 E50833	E50833 Novel G pro
24	307.2	25.9	1002	9 AF406692	AF406692 Homo sapi
25	307.2	25.9	1002	9 AF411113	AF411113 Homo sapi
26	307.2	25.9	1484	6 AX128499	AX128499 Sequence
27	307.2	25.9	2857	9 AF295368	AF295368 Homo sapi
28	305.6	25.7	1002	9 AB083597	AB083597 Homo sapi
29	305.6	25.7	1124	6 AR137021	AR137021 Sequence
30	305.6	25.7	1124	6 AR200987	AR200987 Sequence
31	305.6	25.7	1124	6 E36078	E36078 cDNA clone
32	303.2	25.5	1958	9 AB056816	AB056816 Macaca fa
33	302.4	25.5	1488	6 AR075038	AR075038 Sequence
34	295.4	24.9	1474	6 AR092425	AR092425 Sequence
35	263.2	22.2	356	6 AX369358	AX369358 Sequence
36	252.4	21.3	2416	6 AX329578	AX329578 Sequence
37	252.4	21.3	2416	9 HUMRSC338	D13626 Human mRNA
38	252.4	21.3	8565	9 AF411849S2	AF456925 Homo sapi
39	252.4	21.3	183861	9 AC078816	AC078816 Homo sapi
40	250.8	21.1	1027	6 AR154244	AR154244 Sequence
41	236	19.9	190799	2 AC128510	AC128510 Rattus no
42	230.8	19.4	1785	10 BC028995	BC028995 Mus muscu
43	230.8	19.4	1847	10 AF177211	AF177211 Mus muscu
44	230.8	19.4	191305	2 AC115919	AC115919 Mus muscu
45	219.2	18.5	1690	10 RN076206	U76206 Rattus norv

ALIGNMENTS

```
RESULT 1
BC017898
LOCUS
DEFINITION Homo sapiens, Purinergic receptor P2Y, G protein-coupled, 12, clone
MGC:23802 IMAGE:4251263, mRNA, complete cds.
ACCESSION BC017898
VERSION BC017898.1 GI:17389766
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1424)
AUTHORS Strausberg,R.
TITLE Direct Submission
```

JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) medpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 35 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 1232482.

FEATURES

Source

Location/Qualifiers
1..1424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="LOCUSID:64805"
/clone="MGC:23802 IMAGE:4251263"
/tissue_type="Prostate"
/clone_lib="NH_MGC_83"
/lab_host="DH108"
/note="Vector: pDNR-LIB"
245..1273
/codon_start=1
/product="Purification receptor P2Y, G protein-coupled, 12"
/protein_id="AAH17898.1"
/db_xref="GI:1738976"
/translation="MOAVDNLTSAPGNTSLCTDYKIQVLPILVLPVGLITNG
LAMIPEIORSKNFIILFNKNTISDLITLTPPEKILSDAKITGTPLEPCQVTSY
IFETMYTISLSLGLTIDRYOKTTPRFTSNKRLILGKILSVTMAWFLSLPDM
ILNRPDRKNKCSFLSEFGLWHELVNICOVTEWNLIVCYTLITKELR
SYVTRGVGKVRKVNVRVFIILAVFICFVPHFARIPVTLISQTRVDYEDCAEML
FYVKESTLWLSLMLCLDPIFYFLCKSPRNSLISMLKPNSTSLSDNRKKEQDGG
DPNEETPM"

CDS

BASE COUNT 465 a 309 c 244 g 406 t
ORIGIN
Query Match 97.7%; Score 1160; DB 9; Length 1424;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

28 CTCGAGGCTGCATTAATCTACTACTGATGATTCATTCAACCTCCAGAAATGACAGTT 87
166 CTCGAGGCTGCATTAATCTACTACTGATGATTCATTCAACCTCCAGAAATGACAGTT 225
88 ATCAGGTAACCAAGAAATGCAAGCCGTCGCAATCTCACCCTGCGCTGGAAAC 147
226 ATCAGGTAACCAAGAAATGCAAGCCGTCGCAATCTCACCCTGCGCTGGAAAC 285
148 CAGCTGTGACCAAGACTACAAATTCACCCAGGTCCTCTCCACTGCTTACACTGT 207
286 CAGCTGTGACCAAGACTACAAATTCACCCAGGTCCTCTCCACTGCTTACACTGT 345
208 CCGTGTGTTTGGTGAATTCACAAATGCGCCGCGAGAGAGATTTCTTCAAAATCG 267
346 CCGTGTGTTTGGTGAATTCACAAATGCGCCGCGAGAGAGATTTCTTCAAAATCG 405
268 GAGTAATCAAACTTATTTCTTAAAGACACAGCATTTGATCTCTCAATGAT 327
406 GAGTAATCAAACTTATTTCTTAAAGACACAGCATTTGATCTCTCAATGAT 465
328 TGTGACTTTTCCATTCAAAATCTTATGATGCAACTGGAAACGACCACTGGAAC 387

466 TGTGACTTTTCCATTCAAAATCTTATGATGATGCCAAACTGGAAACAGACCACTGGAAC 525
388 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTATACATGATTTTCATT 447
526 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTATACATGATTTTCATT 585
448 CCTGGAGCTGATTAATCTATGATGCTCTACAGAGACCAAGCCATTTAAACATCCAA 507
586 CCTGGAGCTGATTAATCTATGATGCTCTACAGAGACCAAGCCATTTAAACATCCAA 645
508 CCCCCAAAATCTCTGGGGGCTAAGATTTCTCTGTTGATCTGGGCAATTCATGCTCT 567
646 CCCCCAAAATCTCTGGGGGCTAAGATTTCTCTGTTGATCTGGGCAATTCATGCTCT 705
568 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCGAGACCAAGATGGAAGAA 627
706 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCGAGACCAAGATGGAAGAA 765
628 ATGCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCGAGACCAAGATGGAAGAA 687
766 ATGCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCGAGACCAAGATGGAAGAA 825
688 TCAAGTCAATTTCTGATTAATTTCTTAATTTGTTATGTTATGTTATACATCAATACAA 747
826 TCAAGTCAATTTCTGATTAATTTCTTAATTTGTTATGTTATGTTATACATCAATACAA 885
748 AGAATCTACCGGTGATACATGATTAAGAACAGAGGGGTGATGATTAAGTCCCAAGAAAGT 807
886 AGAATCTACCGGTGATACATGATTAAGAACAGAGGGGTGATGATTAAGTCCCAAGAAAGT 945
808 GAACGTCAAGTCTTCAATTAATGATGCTGATTTCTTATTTGTTTGTCTCTTCCATTT 867
946 GAACGTCAAGTCTTCAATTAATGATGCTGATTTCTTATTTGTTTGTCTCTTCCATTT 1005
868 TGCCCGAATTCCTTACACCTGACCCCAACCCGAGATGTTTGGACGCGCATGCGTGA 927
1006 TGCCCGAATTCCTTACACCTGACCCCAACCCGAGATGTTTGGACGCGCATGCGTGA 1065
928 TACTCTGTATATGTAAGAGAGACACTGTGTGTTAACTTCTTAATATGATGCTGTA 987
1066 TACTCTGTATATGTAAGAGAGACACTGTGTGTTAACTTCTTAATATGATGCTGTA 1125
988 TCGGTTCATGATTTTCTCTTGGCAAGTCCCTCAGAAATCTCTGATATGATGCTGAA 1047
1126 TCGGTTCATGATTTTCTCTTGGCAAGTCCCTCAGAAATCTCTGATATGATGCTGAA 1185
1048 GTGCCCAATTCGCAACATCTCTGCCAGAGCAATGAGAAAGAACAGATGCTGG 1107
1186 GTGCCCAATTCGCAACATCTCTGCCAGAGCAATGAGAAAGAACAGATGCTGG 1245
1108 TGACCCAAATGAAAGAGACTGTAATGTAACAAATTAAGTAAATTTCAATCTTT 1167
1246 TGACCCAAATGAAAGAGACTGTAATGTAACAAATTAAGTAAATTTCAATCTTT 1305
1168 TGTGTTCAAGATCTGTTAA 1187
1306 TGTGTTCAAGATCTGTTAA 1325

RESULT 2
AF313449 1332 bp mRNA linear PRI 12-JAN-2001
LOCUS
DEFINITION Homo sapiens P2Y12 platelet ADP receptor mRNA, complete cds.
ACCESSION AF313449
VERSION AF313449.1 GI:12083901
KEYWORDS

SOURCE

Homo sapiens.
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1332)
Hollinger, G., Jantzen, H.-M., Vincent, D., Li, G., England, L.,
Ramakrishnan, V., Yang, R.-B., Nurdien, P., Nurdien, A., Julius, D. J. and


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
72. .1100
/codon_start=1
/product="G-protein coupled receptor SP199"
/protein_id="AAK00648.1"
/db_xref="GI:12656598"
/translation="MQAVNDLTSAPGNTSLCTRDYKITYQVLEPLLTVLVEVGLITNG
LAMIIFQIQRSSNFIIIFLKNVYISDLMILTFPEKILSDKLSTGTGPRTCVQTSV
IEFTVTVIISISGLGITIDRQKTRTPKSPNSNLGAKILSVIVAMFMILSLPNP
ILINRQPRDKNVKCSFLKSEGLVMEHIVNYICQVLFMIFLVIYCYTILITELIR
SYATRGVGKVPKRKAVKVFIIIAVEFICVPHFAKIPYTSQTRDVECTAEKNTI
FYAKSESTFLMILSLNACIDPFIYFPLCKSFNSMLKCPNSATSLQSDRVEKKEQDGG
DPNEETPM"

```

Query Match Similarity	97.1%	Score 1152;	DB 9;	Length 1152;
Best Local Similarity	100.0%	Pred. No. 1.9e-272;		
Matches 1152;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	36	TGCATATACCTACTCTTCTGATGATCATTTCAAAACCTCCAGAAATCAACAGTTATAGGTA	95
Db	1	TGCATATACCTACTCTTCTGATGATGATTCATTTCAAAACCTCCAGAAATCAACAGTTATAGGTA	60
QY	96	ACCAACAGAAATGCAACCGCGTCGACATTCACCTCTGCGCTGGGAACACCAATCTCT	155
Db	61	ACCAACAGAAATGCAACCGCGTCGACATTCACCTCTGCGCTGGGAACACCAATCTCT	120

QY	156	GCACGAGAGACTACAAAATATCAACCCAGAGTCCCTTTCCCACTGGCTATACACTGTGCTCCGTTT	215
Db	121	GCACGAGAGACTACAAAATATCAACCCAGAGTCCCTTTCCCACTGGCTATACACTGTGCTCCGTTT	180
QY	216	TTGTGTGACTTATCAAAAATGGCTGGCGGATGAGAGATTTTCTTTCAAAATCCGGAGTAAT	275
Db	181	TTGTGTGACTTATCAAAAATGGCTGGCGGATGAGAGATTTTCTTTCAAAATCCGGAGTAAT	240

OY	276	CAACCTTTAATTAATTTTCTTAAAGAACACAGACATTTTCGATCTCTCCAGATTCGACATT	335
Db	241	CAAACTTTATTTATTTTCTTAAAGAACACAGCATTTTCGATCTTCTCATGATTCGACTT	300
OY	336	TTTCATTTCAAAAATTTCTAGTAGTGAAGCCAAATCTGGGAACAGACACACTGGAACATTTTGTCT	395
Db	301	TTTCATTTCAAAAATTTCTAGTAGTGAAGCCAAATCTGGGAACAGACACACTGGAACATTTTGTCT	360

Accession	Sequence	Position
QY	GTCGGTAACCTCGCTAATTTATTTACAAATATATCATGATTTTCATTCCCTGGGAC	455
Db	GTCAGTTACCTCGCTCAATTTTATTTTTCACAACTATATCATGATTTTCATTCCCTGGGAC	420
QY	TGATTAACATGATCGCTACCAAGACACACAGGCCATTATTAACATCATCAACCCCAAAA	515
Db	TGATTAACATGATCGCTACCAAGACACACAGGCCATTATTAACATCATCAACCCCAAAA	480

QY	516	ATCTCTTTGGGGGCTAAGACTTCTCTCTGTGTTGTCATCTGGGCAATCAATGTCCTTACTCTCTT	575
Db	481	ATCTCTTTGGGGGCTAAGAGTTCTCTCTGTGTTGTCATCTGGGCAATCAATGTCCTTACTCTCTT	540
QY	576	TGCCATCATCATTTCTGACCAACAGGCGACCGAGAGACAAGAAATGTGTGAAGAAATGCTCTT	635
Db	541	TGCCATCATCATTTCTGACCAACAGGCGACCGAGAGACAAGAAATGTGTGAAGAAATGCTCTT	600

QY	636	TCCCTTAAATCAGAGTTCGGCTAGTCTGGCAGTAAGTAAGTAAATTCATCTGTCAAGTCA	695
Db	601	TCCCTTAAATCAGAGTTCGGCTAGTCTGGCAGTAAGTAAGTAAATTCATCTGTCAAGTCA	660
QY	696	TTTTCTGATTAATTTCTTAATGTATGTATGTATACACTCTATPACAAAGACGCT	755
Db	661	TTTTCTGATTAATTTCTTAATGTATGTATGTATACACTCTATPACAAAGACGCT	720

OY		756	ACCGGTCATACGTAGAAGCAAGGGGCTGTAGGTAAAGTCCCGAGGAAGTTGAACGTCA	815
Dδ		721	ACCGETCATACGTAAAGAAGGAGGGGTGTAGGTAAGTCCCAGCAGAAAAAAGGTGAACGTCA	780

Oy	816	AAGTTTCATATATATCATTGCTGTGATTCCTTAAATGTTGGTTCCTTGCATTTGGCCGAA	875
Db	781	AAGTTTCAITATATCATTTCTGCTGATCTTATTTGTTGTTGTTCTTCATTTTGGCCGAA	840
Oy	876	TTCCCTTACACCCCTTGAGCCAAACCCGGGATGTCCTTTGACATGCACTGCTCAATAATACTGT	935
Db	841	TTCCCTTACACCCCTTGAGCCAAACCCGGGATGTCCTTTGACATGCACTGCTCAATAATACTGT	900
Oy	936	TCTATGTAAAGAGAGACACTGTGTGTTAACTTCTTAAATGACATGCTGGATCCGTTC	995
Db	901	TCTATGTAAAGAGAGACACTGTGTGTTAACTTCTTAAATGACATGCTGGATCCGTTC	960
Oy	996	TCTATTTTTCCTTTGGCAAGTCCTTCAGAAATTCCTGATATAGATGTGTAAGTCCCA	1055
Db	961	TCTATTTTTCCTTTGGCAAGTCCTTCAGAAATTCCTGATATAGATGTGTAAGTCCCA	1020
Oy	1056	ATTTCGCAACATCTCTGTGCCAGGACCAATPAGGAAAAAAGAACAGAGATGTGTGACCCAA	1115
Db	1021	ATTTCGCAACATCTCTGTGCCAGGACCAATPAGGAAAAAAGAACAGAGATGTGTGACCCAA	1080
Oy	1116	ATGAAGAGACTCCAAATGTAAACAATTAAGTAAAGAAATATTCAATCTTTTGTGTCA	1175
Db	1081	ATGAAGAGACTCCAAATGTAAACAATTAAGTAAAGAAATATTCAATCTTTTGTGTCA	1140
Oy	1176	GAACTCGTTTAA	1187
Db	1141	GAACTCGTTTAA	1152

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AB056385	AB056385	1951 bp	mRNA	linear	PRI 01-MAR-2001	
		Macaca fascicularis brain cDNA, clone:OflA-10912.				
	AB056385					
	AB056385.1	GI:1358631				
	fis (full insert sequence).					
	Macaca fascicularis adult male frontal lobe left cDNA to mRNA.					

ORGANISM	REFERENCE
Chong-Li: macaque brain cDNA library with clone: Q11A-10912.	
Macaca fascicularis	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;	
Cercopithecinae; Macaca.	
1 (sites)	
Ossada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,	
Terao, K., Sugaki, Y., Sugano, S. and Hashimoto, K.	
Isolation of full-length cDNA clones from macaque brain cDNA	
library	

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1951)
AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genbank/, Tel: 81-3-5285-1111 (ext. 2120), Fax: 81-3-5285-1181)

Vector: pME18S-Fl3 (Acc. No. AB009864)
R. Site1: DraIII (CACTGTGTC)
R. Site2: DraIII (CACCATGTGC)
Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCGTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-flg3. XhoI sites just outside

FEATURES
SOURCE
Location/Qualifiers
1. .1951

	Query Match	Best Local Similarity	97.3%	Score 1110.4	DB 9	Length 1951	
	Matches 1129	Conservative	0	Mismatches 31	Indels 0	Gaps 0	
BASE COUNT	658 a	395 c	308 g	590 t			
ORIGIN							
QY	28	CTGCAGCTGCAATTAAGTACTACTTACCTGATTCAGATTCACCAACCTTCAGAACTCAACAGTT	87				
DB	169	CTGCAGCTTGCATTAATTAAGTACTACTTACCTGATTCAGATTCACCAACCTTCAGAACTCAACAGTT	228				
OY	88	ATTCAGCTAACCAAGAAATGCAAGCCGTGACAACTCACCTCCTGGCCCTGGGAAC	147				
DB	229	ATTCAGCTAACCAAGAAATGCAAGCCGTGACAACTCACCTCCTGGCCCTGGGAAC	288				
OY	148	CAGTCTGTGCACACGAGACTACAAATATCACCCAGGCTCTTCCACATGCTCTACACTGT	207				
DB	289	CAGTCTGTGCACACGAGACTACAAATATCACCCAGGCTCTTCCACATGCTCTACACTGT	348				
OY	208	CCTGTTTTTGTGTGGACTTATCACAAATGCGCTGGCGATGAGAGATTTTCTTCAAAATCCG	267				
DB	349	CCTGTTTTTGTGTGGACTTATCACAAATGCGCTGGCGATGAGAGATTTTCTTCAAAATCCG	408				
OY	268	GAGTAAATCAAACTTATTTTCTTAAAGACACATCTATTTCTGATCTTCTCATGAT	327				
DB	409	GAGTAAATCAAACTTATTTTCTTAAAGACACATCTATTTCTGATCTTCTCATGAT	468				
OY	328	TCTGCATTTTCATTTCAAAATCTTATGATGATGCCAAATGGGAACAGACCACTGAGAAC	387				
DB	469	TCTGCATTTTCATTTCAAAATCTTATGATGATGCCAAATGGGAACAGACCACTGAGAAC	528				
OY	388	TTTTGTGTGTCAAGTTACCTCGCTCATATTTTATTTTCACAATGTAATACGATTTTCAT	447				
DB	529	TTTTGTGTGTCAAGTTACCTCGCTCATATTTTATTTTCACAATGTAATACGATTTTCAT	588				
OY	448	CCTGGGACTGATTAATATCGATCGTTACCAAGAACCCAGGCCATTTAAACATCCAA	507				
DB	569	CCTGGGACTGATTAATATCGATCGTTACCAAGAACCCAGGCCATTTAAACATCCAA	648				
OY	508	CCCCAAAATATCTCTGGGGGCTAAGATTCCTGCTGTGTCATCTGGGATTCATGTTCTT	567				
DB	649	CCCCAAAATATCTCTGGGGGCTAAGATTCCTGCTGTGTCATCTGGGATTCATGTTCTT	708				
OY	568	ACTCTCTTTGGCTTAACATGATTTCTGACCAACAGGCGAGCGAGAGACAAGAAATGTGAAGAA	627				
DB	709	ACTCTCTTTGGCTTAACATGATTTCTGACCAACAGGCGAGCGAGAGACAAGAAATGTGAAGAA	768				
OY	628	ATGCTCTTTCTTAATATAGAGTTGCGTCTAGTCTGGCATGAATATGTAATTTACATCTG	687				
DB	769	ATGCTCTTTCTTAATATAGAGTTGCGTCTAGTCTGGCATGAATATGTAATTTACATCTG	828				
OY	688	TCAAGTCAATTTCTGATTAATATTTCTTAATGTTTATGATGTTTATACACTATTAACAA	747				
DB	829	TCAAGTCAATTTCTGATTAATATTTCTTAATGTTTATGATGTTTATACACTATTAACAA	888				

OY	748	AGAACTGACGGGTCAATCGTAAGAAGAGGGGGTGTAGCTAAGATCCCCAGAGAAAAAGGT	807
Db	889	AGAACTGACGGGTCAATCGTAAGAAGAGGGGGTGTAGCTAAGATCCCCAGAGAAAAAGGT	948
OY	808	GNAAGTCAAAAGTTTTCATTATCATATGCAGTAATCATTATTTGTTTTGTTGCCATTTT	867
Db	949	GAGCTCAAAGTTTTCATTATCATATGCAGTAATCATTATTTGTTTTGTTGCCATTTT	1008
OY	868	TGCCCGAATTCCTPACACCCTGAGCCAAACCCGGGAGTGCTTTTGACTGACATGCTGA AAA	927
Db	1009	TGCCCGAATTCCTPACACCCTGAGCCAAACCCGGGAGTGCTTTTGACTGACATGCTGA AAA	1068
OY	928	TACTCTGTTCTAATGGAAGAAGACACATCTGTGTAACTCTCTTAATTCATCCTGTGA	987
Db	1069	TACTCTGTTCTAATGGAAGAAGACACATCTGTGTAACTCTCTTAATTCATCCTGTGA	1128
OY	988	TCCGTTTCATTTATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTGTGAATATGCTGAA	1047
Db	1129	TCCGTTTCATTTATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTGTGAATATGCTGAA	1188
OY	1048	GTGCCCCAATTCGTGCACATCTCTGTGCCAGACAAATRGAAAAAACAAGCATGTGG	1107
Dc	1189	GTGCCCCAATTCGTGCACATCTCTGTGCCAGACAAATRGAAAAAACAAGCATGTGG	1248
OY	1108	TGACCCAAATGGAAGAGATCCAAATGTAAACAATTAATAAGAAATTTTCATTCCTT	1167
Dd	1249	TGACCCAAATGGAAGAGATCCAAATGTAAACAATTAATAAGAAATTTTCATTCCTT	1308
OY	1168	TGTGTTTCAGAACCTGTTAAA	1187
Dd	1309	TGCGTTTCAGAACCTGTTAAA	1328
RESULT 5			
AB062981			
LOCUS			
DEFINITION	AB062981	2266 bp mRNA linear PRI 13-JUN-2001	
ACCESSION	AB062981	Macaque fascicularis brain cDNA clone:QmoA-10634, full insert sequence.	
VERSION	AB062981.1	GI:14388396	
KEYWORDS		oligo capping; fts (full insert sequence).	
SOURCE		Macaque fascicularis adult male medulla oblongata cDNA to mRNA, clone_1fb:macaque brain cDNA library QmoA clone:QmoA-10634.	
ORGANISM		Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.	
REFERENCE		1 (sites)	
AUTHORS		Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M., Terao,K., Shidjuku-Ku, Tokyo 162-8640, Japan	
TITLE		Isolation of full-length cDNA clones from macaque brain cDNA libraries	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 2266)	
AUTHORS		Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.	
TITLE		Direct Submission	
JOURNAL		Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-Ku, Tokyo 162-8640, Japan	
		(E-mail:khashih@nii.go.jp, URL:http://www.nii.go.jp/yoken/genebank/, Tel:-81-3-5285-1111(ex 2120), Fax:-81-3-5285-1181)	
COMMENT		Lab host: TOP10	
		Vector: pME18S-FU3 (ACC NO. AB009864)	
		R. Site1: DraIII (CACCTGTGCG)	
		R. Site2: DraIII (CACCATGTG)	
		Description: 1st strand cDNA was primed with an oligo(dt2) primer [AATGTCGCTTTTATTTTATTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb.The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FU3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries	

were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).
Custom primer used for sequencing (5' end primer (CTGTGCTCTAAAGCTGG): 3' end primer (CGACCTGACCTGACGACAC)).
Location/Qualifiers

FEATURES

Source

1. 2266
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="OmoA-10634"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="macaque brain cDNA library OmoA"
/dev_stage="adult"
240. 1268
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB60747.1"
/db_xref="GI:14388397"
/translation="MQAIDNLTSAPGNTSLCTRDYKITGVLPFLYTVLEFGLITNS
LARIFFOIRKSNFIIFLKNVSDILMLITPPLILSPAKIAGAPLEFVQVTSV
IEFTYISISGLITIDRYOKTRPESKSNKLLGKILSLVILMAEFLIPLNM
ILNRRPDKNNKCSFLKSEGLVMEHLYNYTCOVLPINPLIVTCYTLITKELYR
SVYRPGVGVKPKRKNVYFIIIVAFICFPVPHARIPYITISOTRDYFDCAENTL
FYKESTLMLTSLNACLDPETIFFLCKSEFNSLISMLKCPNSATSOODNRKEDQGG
DPNEETPM"

CDS

BASE COUNT 794 a 431 c 349 g 692 t
ORIGIN

Query Match

Best Local Similarity 97.28; Score 1108.8; DB 9; Length 2266;
Pred. No. 7; 7e-262;

Matches 1128; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 28 CTGACGAGCTCAATTAATCTACTACTGATGATATCAATCAACCTCCAGATCAACAGTT 87
DB 161 CTGACGCTTGCANAACTACTATTTACTGATATCAATCAATCCCTCCAGATCAACGTT 220
QY 88 ATCAGGTACCAACAGAAATGCAAGCGCTGACATCTCACTGCGCTGGGAACAC 147
DB 221 ATCAGGTACCAACAGAAATGCAAGCGCTGACATCTCACTGCGCTGGGAACAC 280
QY 148 CAGCTGTCGACAGACAGACAGAAATGCAAGCGCTGACATCTCACTGCGCTGGGAACAC 207
DB 281 CAGCTGTCGACAGACAGACAGAAATGCAAGCGCTGACATCTCACTGCGCTGGGAACAC 340
QY 208 CCGTGTGTTGTTGACTTATCAAAATGCGCGGATGAGATTTCTTTCAATCCG 267
DB 341 CCGTGTGTTGTTGACTTATCAAAATGCGCGGATGAGATTTCTTTCAATCCG 400
QY 268 GAGTAATCAAACTTATTTCTTAAGAACACAGTATTTGATCTTCATGAT 327
DB 401 GAGTAATCAAACTTATTTCTTAAGAACACAGTATTTGATCTTCATGAT 460
QY 328 TGTGACTTTCATCAAAATCTTACTGATGCAAACTGGGAACAGACACACGAGAAC 387
DB 461 TGTGACTTTCATCAAAATCTTACTGATGCAAACTGGGAACAGACACACGAGAAC 520
QY 388 TTTTGTGTTCAAGTACCTCCGATATTTTATTTTCAATGATATAGTATTTGAT 447
DB 521 TTTTGTGTTCAAGTACCTCCGATATTTTATTTTCAATGATATAGTATTTGAT 580
QY 448 CTTGGGACGTATACTATGATGCTACGAGAACAGACAGGACGATTTAAACATCAA 507
DB 581 CTTGGGACGTATACTATGATGCTACGAGAACAGACAGGACGATTTAAACATCAA 640
QY 508 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTGTTGATCTGGGACATTCATGTTCTT 567
DB 641 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTGTTGATCTGGGACATTCATGTTCTT 700
QY 568 ACTCTCTTCTTACATGATTTGACCAACAGGACGAGAGACAAAGATGGAAGAA 627
DB 701 ACTCTCTTCTTACATGATTTGACCAACAGGACGAGAGACAAAGATGGAAGAA 760

QY 628 ATGCTCTTCCCTTAATACAGATTCGCTAGTCTGCGACAAATATTAATATCATCTG 687
DB 761 ATGCTCTTCCCTTAATACAGATTCGCTAGTCTGCGACAAATATTAATATCATCTG 820
QY 688 TCAAGTATTTTTCGATTAATTTCTTAATTTGATTTGATTAATACATCATTAACA 747
DB 821 TCAAGTATTTTTCGATTAATTTCTTAATTTGATTTGATTAATACATCATTAACA 880
QY 748 AGAAGCTACCGGTCATACGTAAGAACAGAGGGGTGATGAAGTCCCGAGGAAAAGGT 807
DB 881 AGAAGCTACCGGTCATACGTAAGAACAGAGGGGTGATGAAGTCCCGAGGAAAAGGT 940
QY 808 GAAAGTCAAAATTTTCATTAATTCATTTGCTGTAATTTGATTTGCTTTTCCATTT 867
DB 941 GAAAGTCAAAATTTTCATTAATTCATTTGCTGTAATTTGATTTGCTTTTCCATTT 1000
QY 868 TGCCCGAATTCCTTACACCTCGAGCCAAACCCGGAGATGCTTTGACTGCACTGTA 927
DB 1001 TGCCCGAATTCCTTACACCTCGAGCCAAACCCGGAGATGCTTTGACTGCACTGTA 1060
QY 928 TACTCTGTGTGTAAGAGAGACACTGCTGTTAATCTTAAATGATGCTGGA 987
DB 1061 TACTCTGTGTGTAAGAGAGACTGCTGTTAATCTTAAATGATGCTGGA 1120
QY 988 TCCGTTATCTATTTTTCCTTTGCAAGTCCCTGAGAAATCTTGTATGATGCTGAA 1047
DB 1121 TCCGTTATCTATTTTTCCTTTGCAAGTCCCTGAGAAATCTTGTATGATGCTGAA 1180
QY 1048 GTGCCCCAATTTGCAACATCTCTGTCGACGACATAGAAAAAGACAGATGCTG 1107
DB 1181 GTGCCCCAATTTGCAACATCTCTGTCGACGACATAGAAAAAGACAGATGCTG 1240
QY 1108 TGACCCCAATGAGAGACTCCATGTAACAAATTAAGGAAATTTCAATCTT 1167
DB 1241 TGACCCCAATGAGAGACTCCATGTAACAAATTAAGGAAATTTCAATCTT 1300
QY 1168 TGTGTTGAGACTGTTAAA 1187
DB 1301 TGTGTTGAGACTGTTAAA 1320

RESULT 6
AC117393/c 328187 bp DNA linear HTG 29-APR-2002
LOCUS
DEFINITION
Homo sapiens chromosome 3 clone Rpl1-372J9, *** SEQUENCING IN
PROGRESS ***, 17 unordered pieces.
AC117393 AC012057
AC117393.2 GI:20279204
VERSION
HTG: HTGS_PHASE1; HTGS_ACTIVEFIN.
KEYWORDS
Homo sapiens.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 328187)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alshrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,R., Cavazos,S.R., Chacko,J., Chaves,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earmhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, V., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.,
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Lonsdale, H.,
Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mel, G., Metker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabath, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, S., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokkenko, S., Ogub, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L. L., Qulles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scheerer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Syatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 328187)
Worley, K. C.

Direct Submission
Submitted (10-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 328187)
Worley, K. C.

Direct Submission
Submitted (29-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On or before Apr 26, 2002 this sequence version replaced
gi:8072477, gi:20127689.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HDGP
Center clone name: RP11-372J9
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 345031 bases at least Q40
Consensus quality: 361075 bases at least Q30
Consensus quality: 371672 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
2010: contig of 2010 bp in length
*
* 2011 2110: gap of unknown length
*
* 2111 4550: contig of 2440 bp in length
*
* 4551 4650: gap of unknown length
*
* 4651 7410: contig of 2660 bp in length
*
* 7411 7410: gap of unknown length
*
* 7411 10060: contig of 2650 bp in length
*
* 10061 10160: gap of unknown length
*
* 10161 14011: contig of 3851 bp in length

14012 14111: gap of unknown length
*
* 14112 18735: contig of 4624 bp in length
*
* 18736 18835: gap of unknown length
*
* 18836 23973: contig of 5138 bp in length
*
* 23974 24074: gap of unknown length
*
* 24074 30069: contig of 5996 bp in length
*
* 30070 30169: gap of unknown length
*
* 30170 37639: contig of 7470 bp in length
*
* 37640 37739: gap of unknown length
*
* 37740 46044: contig of 8305 bp in length
*
* 46045 46144: gap of unknown length
*
* 46145 57813: contig of 11659 bp in length
*
* 57814 57913: gap of unknown length
*
* 57914 73711: contig of 15798 bp in length
*
* 73712 73812: gap of unknown length
*
* 73812 89444: contig of 15633 bp in length
*
* 89445 89544: gap of unknown length
*
* 89545 110717: contig of 21173 bp in length
*
* 110718 110817: gap of unknown length
*
* 110818 131687: contig of 20870 bp in length
*
* 131688 131787: gap of unknown length
*
* 131788 165404: contig of 33617 bp in length
*
* 165405 165505: gap of unknown length
*
* 165505 328187: contig of 162683 bp in length.

FEATURES
source
1. 328187
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-372J9"

BASE COUNT 98046 a 61482 c 63090 g 103950 t 1619 others

ORIGIN

Query Match 92.4% Score 1097; DB 2: Length 328187;
Best Local Similarity 100.0%; Pred. No. 5.7e-259;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 AGGTACCAACAAAGAAAGCAAGCCGTGACATCTTCACCTTGGCCGTGGACACACG 150
Db 268734 AGGTACCAACAAAGAAAGCAAGCCGTGACATCTTCACCTTGGCCGTGGACACACG 268675

OY 151 TCTGTGACACAGACACTCAAAATACCCAGGCTCTTCCACATGCTCTACATGTCCT 210
Db 268674 TCTGTGACACAGACACTCAAAATACCCAGGCTCTTCCACATGCTCTACATGTCCT 268615

OY 211 GTTTTGTGGACTTATCACAAATGGCTGGCATGAGATTTCTTCAATCCGAG 270
Db 268614 GTTTTGTGGACTTATCACAAATGGCTGGCATGAGATTTCTTCAATCCGAG 268555

OY 271 TAAATCAAACTTATTTATTTCTTAAAGACACTATTTCTGATCTTCATGATTC 330
Db 268554 TAAATCAAACTTATTTATTTCTTAAAGACACTATTTCTGATCTTCATGATTC 268495

OY 331 GACTTTTCATTCAAAATTTCTAGATGCCAAACTGGGACAGACACATGACATCTT 390
Db 268494 GACTTTTCATTCAAAATTTCTAGATGCCAAACTGGGACAGACACATGACATCTT 268435

OY 391 TGTGTGCAAGTGTACCTCGCATATTTTATTTTCAATGTATTCAGATTTCAATTCCT 450
Db 268434 TGTGTGCAAGTGTACCTCGCATATTTTATTTTCAATGTATTCAGATTTCAATTCCT 268375

OY 451 GGGAGCTGAATCATGATCGCTACACAGACACACACGACGATTTAAACATCCACCC 510
Db 268374 GGGAGCTGAATCATGATCGCTACACAGACACACACGACGATTTAAACATCCACCC 268315

OY 511 CAAAATCTCTTTGGGGGCTAGATTTCTCTGTGTCTATCTGGGATTCATGTTTACT 570
Db 268314 CAAAATCTCTTTGGGGGCTAGATTTCTCTGTGTCTATCTGGGATTCATGTTTACT 268255

OY 571 CTCTTGGCTTAACATGATTCGACCAACAGGACGCGAGAGACAGATGGAAGAAAG 630
Db 268254 CTCTTGGCTTAACATGATTCGACCAACAGGACGCGAGAGACAGATGGAAGAAAG 268195

Db 996 CCCCAATTCGCAACATCTCTGTCGCCAGACATAGCAAAAAGACAGATGGTGTGA 1055
QY 1111 CCCAATGAAGAGACCTCCATGTAACCAATTAAGCAATATTTCAATCTTTGT 1170
Db 1056 CCCAATGAAGAGACCTCCATGTAACCAATTAAGCAATATTTCAATCTTTGT 1115
QY 1171 GTTCAGAACCTGTTAA 1187
Db 1116 GTTCAGAACCTGTTAA 1132

RESULT 8
AF310685 1143 bp DNA linear PRI 17-NOV-2001
LOCUS Homo sapiens ADP-glucose receptor gene, complete cds.
DEFINITION AF310685
ACCESSION AF310685
VERSION AF310685.1 GI:16973448
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1143)
AUTHORS Reinscheid, R.K., Nothacker, H.-P., Wang, Z., Zeng, J., Ehler, F.J. and
Civelli, O.
TITLE ADP-glucose activates a G-protein coupled receptor and inhibits
smooth muscle contractions
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1143)
AUTHORS Nothacker, H.-P., Reinscheid, R.K., Wang, Z. and Civelli, O.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2000) Pharmacology, UC Irvine, 354 Med Surge II,
Irvine, CA 92697, USA

FEATURES
source
1. 1143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q21-q25"
/product="ADP-glucose receptor"
/note="ADP-glucose receptor"
/product="ADP-glucose receptor"
/note="ADP-G-R: G protein-coupled receptor"
/codon_start=1
/product="ADP-glucose receptor"
/protein_id="AF310685.1"
/translation="MQAVNDLTSAPGNTSICRDKITTOVLEPLATVLEFVGLING
LAKRIFQIRSKSNFIIFLKNYISDLMLITPEKILSDAKLGTGRLTFVCOVTSV
IFEFWYISISFLGILITIDYQKOTTRPFKTSNPKNLGAILSVIWAFFLLSLPM
ILNRPDRKNSKCSFLKSEFLVWHEIVNYICQVFWINFLIVICVYLITKEIYR
SVYRTGCVGVPRKRVNRYVFIILAVFICFVPHFARIPIYLSOTRDVDCAEANTL
FYVKESTLMTLSLNACLDPIYFELFKSPFNSLSMLKCNKNSATLSLSDNRKKEQGG
DPRERTPM"

BASE COUNT 326 a 257 c 201 g 359 t
ORIGIN

Query Match 92.1%; Score 1093.8; DB 9; Length 1143;
Best Local Similarity 99.8%; Pred. No. 3.8e-258;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGTAAACCAAGAAATGCAAGCCGTCGACATCTGACCTGCGCTGGGAACACG 150
Db 36 AGTAAACCAAGAAATGCAAGCCGTCGACATCTGACCTGCGCTGGTAACACG 95
QY 151 TCTGTGACACAGACATCAAAATACCCAGGCTCTTCCCACTGCTACACTGTCT 210
Db 96 TCTGTGACACAGACATCAAAATACCCAGGCTCTTCCCACTGCTACACTGTCT 155
QY 211 GTTTTGTGACACTTATCAAAATGCGCTGGCATGAGATTTCTTTCAATCCGAG 270
Db 156 GTTTTGTGACACTTATCAAAATGCGCTGGCATGAGATTTCTTTCAATCCGAG 215

QY 271 TAAATCAACTTTATTTTCTTAAAGACACATATTTGATCTTCTCAAGTCT 330
Db 216 TAAATCAACTTTATTTTCTTAAAGACACATATTTGATCTTCTCAAGTCT 275
QY 331 GACTTTCACATCAAAATCTTACTGATGCCAATCGGAACAGACACATGAACTTT 390
Db 276 GACTTTCACATCAAAATCTTACTGATGCCAATCGGAACAGACACATGAACTTT 335
QY 391 TGTGTGCAAGTTCACCTCCGTCATATTTTATTTTCAATGATATGATTTCTTCT 450
Db 336 TGTGTGCAAGTTCACCTCCGTCATATTTTATTTTCAATGATATGATTTCTTCT 395
QY 451 GGCAGTATTAATCTATCGATCGCTACCAAGACACAGCCATTTAAACATCAACC 510
Db 396 GGCAGTATTAATCTATCGATCGCTACCAAGACACAGCCATTTAAACATCAACC 455
QY 511 CAAAATCTCTGGGGGCTAAGATTTCTGTTGTATCTGGGCATTCATGTTCTTACT 570
Db 456 CAAAATCTCTGGGGGCTAAGATTTCTGTTGTATCTGGGCATTCATGTTCTTACT 515
QY 571 CTCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAAGATGGAAGAAATG 630
Db 516 CTCTTTCCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAAGATGGAAGAAATG 575
QY 631 CTCTTTCCTTAACATGATTTGACCTAGTCTGATGATTAATGATTAATGATCTGCA 690
Db 576 CTCTTTCCTTAACATGATTTGACCTAGTCTGATGATTAATGATTAATGATCTGCA 635
QY 691 AGTATTTTTCGAGATTAATTTCTTAATTTGATTTGATTTGATTTGATTTGATTTGAT 750
Db 636 AGTATTTTTCGAGATTAATTTCTTAATTTGATTTGATTTGATTTGATTTGATTTGAT 695
QY 751 ACTGTACCGGTCATACGATTAAGACAGAGGGGTGATTAAGTCCCAAGAAAGGTGAA 810
Db 696 ACTGTACCGGTCATACGATTAAGACAGAGGGGTGATTAAGTCCCAAGAAAGGTGAA 755
QY 811 CGTCAAGATTTTCATATATCATTTCTGATTTCTTATTTGTTTCTTCTTCAATTTGC 870
Db 756 CGTCAAGATTTTCATATATCATTTCTGATTTCTTATTTGTTTCTTCTTCAATTTGC 815
QY 871 CCGAATTCCTTACACCCGAGCCAAACCCGGAGTCTTTCACGCACTGCTGAAATAC 930
Db 816 CCGAATTCCTTACACCCGAGCCAAACCCGGAGTCTTTCACGCACTGCTGAAATAC 875
QY 931 TCTGTGTATGTGAAGAGACACTGTGTGATTAATTTCTTAAATGATGCTGATTC 990
Db 876 TCTGTGTATGTGAAGAGACACTGTGTGATTAATTTCTTAAATGATGCTGATTC 935
QY 991 GTTCATCTATTTTCTTCTTGAAGTCTTGAAGATTTCTTGAATGATGCTGAAAGTG 1050
Db 936 GTTCATCTATTTTCTTCTTGAAGTCTTGAAGATTTCTTGAATGATGCTGAAAGTG 995
QY 1051 CCCCAATTCGCAACATCTCTGTCGCCAGCAATAGGAAAAAACAAGATGGTGTGA 1110
Db 996 CCCCAATTCGCAACATCTCTGTCGCCAGCAATAGGAAAAAACAAGATGGTGTGA 1055
QY 1111 CCCAATGAAGAGACCTCCATGTAACCAATTAAGCAATATTTCAATCTTTGT 1170
Db 1056 CCCAATGAAGAGACCTCCATGTAACCAATTAAGCAATATTTCAATCTTTGT 1115
QY 1171 GTTCAGAACCTGTTAA 1187
Db 1116 GTTCAGAACCTGTTAA 1132

RESULT 9
AC024886 173270 bp DNA linear PRI 28-MAR-2002
LOCUS Homo sapiens 3 BAC Rpl1-25K24 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
ACCESSION AC024886
VERSION AC024886.20 GI:16572827
KEYWORDS HTG.

SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 173270)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbieri,J., Benton,J., Bimange,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burich,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garra,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,D., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Meier,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okwuonu,G., Oragunye,L., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Szatk,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 173270)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (02-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 173270)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 4 (bases 1 to 173270)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 1, 2001 this sequence version replaced gi:16519437.
 INFORMATION: http://www.hgsc.bcm.tmc.edu/or_email/gc-helpebcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
 Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

Location/Qualifiers

source

1..173270
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-25K24"
 /complement(259..684)
 /rpt_family="MER57A"
 1209..1372
 /rpt_family="L2"
 complement(12596..2721)
 /rpt_family="Alusq/x"
 2722..2897
 /rpt_family="Alusp"
 2980..3091
 /rpt_family="MIR"
 3781..3979
 /rpt_family="MIR"
 complement(4002..4186)
 /rpt_family="MER5A"
 4476..4950
 /rpt_family="L1MA9"
 7145..7190
 /rpt_family="CATATA)n"
 8285..8320
 /rpt_family="CATATA)n"
 8380..8400
 /rpt_family="AT-rich"
 complement(8489..8520)
 /rpt_family="L2"
 complement(8642..8888)
 /rpt_family="MIR"
 complement(8972..9105)
 /rpt_family="L2"
 10672..11018
 /rpt_family="L1MB7"
 11019..11332
 /rpt_family="Alusp"
 11333..11418
 /rpt_family="L1MB7"
 complement(11419..11707)
 /rpt_family="Alusx"

	Query Match	92.1%	Score 1093.8	DB9:	Length 173270;
	Best Local Similarity	99.8%	Pred. No. 3.5e-258;		
	Matches 1095;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps	
Qy	91	AGGTACCAACAAAGAAATGCAAGCCGTGCACAAATCTGACCTCTGCGCCTGGGAACACACAG	150		
Db	127570	AGGTAAACCAACAAAGAAATGCAAGCCGTGCACAACTCACTCTGCGCCTGGTAAACACAG	1276229		
Qy	151	TCTGGACACGAGACGTACAAAAATCACCCAGGCTCTTCCCACTGGCTACACTGTGCT	210		
Db	127630	TCTGTGACACGAGACGTACAAAAATCACCCAGGCTCTTCCCACTGGCTACACTGTGCT	1276889		
Qy	211	GTCTTTTGTGTGGACTTATCACAAATGCGCTGGCGATGAGATTTTCTTTCAAAATCCGGAG	270		
Db	127690	GTCTTTTGTGTGGACTTATCACAAATGCGCTGGCGATGAGATTTTCTTTCAAAATCCGGAG	127749		
Qy	271	TAAATCAAACTTATTTATTTTCTTAAGAACACAGTATTTCTGATCTTCATGATATCT	330		
Db	127750	TAAATCAAACTTATTTATTTTCTTAAGAACACAGTATTTCTGATCTTCATGATATCT	127809		
Qy	331	GACTTTTCATTCAAAAATCTCTTAAGATGCCAAACTGGGAACAGGACCACTGAGAACTTT	390		
Db	127810	GACTTTTCATTCAAAAATCTCTTAAGATGCCAAACTGGGAACAGGACCACTGAGAACTTT	127869		

LOCUS	AX369349	1029 bp	DNA	linear	PAT 16-FEB-2002
DEFINITION	AX369349	Sequence 1	from Patent WO0202599.		
ACCESSION	AX369349				
VERSION	AX369349.1	GI:18857274			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Wattler,F., Wattler,S., Trommler,P. and Nehls,M.C.				
TITLE	Human g protein-coupled receptor igpcr17, and uses thereof				
JOURNAL	Patent: WO 0202599-A 1 10-JAN-2002;				
	Ingenium Pharmaceuticals AG (DE)				

FEATURES	Location/Qualifiers
SOURCE	1..1029 /organism="Homo sapiens" /db_xref="taxon:9606" 1..1029
CDS	/note="unnamed protein product" /codon_start=1 /protein_18="CND2844.1" /translation="MQAVDNLTSAPGNTSLCTRDYKIVQLVPLLYTVLFEVGLITNG LAMIIFQIRKSNFIIFLKNVVISDLMLILFEPKILSDKLEGTGPIRFEVCQVTSV IEFTYPMISISFELGITIDRKOTTRPRTKSPNLSGKATLSVYNAFPLSLPNM ILTRQPDKRVKRCSPKLSRGLVMEIIVYICQVFMIFELIYICVYLITELIR SYTRRGVGVKVPKKRVNKVFTITLAVFICQVPHPRARIPPTLSQTRDVEDCTAENTIL FYVESTLMTSLNACLDPIYFLFKSFRNLSLMLKCPNSATLSISDNNKRCDDGG DNEBETP"
BASE COUNT	285 a 234 c 186 g 324 t
ORIGIN	
Query Match	86.6% Score 1027.4; DB 6; Length 1029; Best Local Similarity 99.9%; Pred. No. 7.8e-242;
Matches 1028; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	107 ATGCAAGCGGTGACAAATCTCAGCTCTGGCGCTGGGAACACAGCTGTGCACGAGAC 166
Db	1 ATGCAAGCGGTGACAACTCTCCTGTGGCTGGGAACACAGCTGTGCACGAGAC 60
QY	167 TCAAAATCACCAAGCTCTCTCCACGTGCTACAGCTGTCCTTTTGTGGACTT 226
Db	61 TCAAAATCACCAAGCTCTCTCCACGTGCTACAGCTGTCCTTTTGTGGACTT 120
QY	227 ATCACAATGGCTTGGCGATGAGATTTTCTTCAATCCGAGTAATCAACTTTAT 286
Db	121 ATCACAATGGCTTGGCGATGAGATTTTCTTCAATCCGAGTAATCAACTTTAT 180
QY	287 ATTTTCTTAAGACACAGTCATTTCTGATCTCTCATGATCTGACATTTCCATTCAA 346
Db	181 ATTTTCTTAAGACACAGTCATTTCTGATCTCTCATGATCTGACATTTCCATTCAA 240
QY	347 ATTTTCTTGTAGTCCAACTGGGAACAGACACACAGTCTTGTGTGCAATTTACC 406
Db	241 ATTTTCTTGTAGTCCAACTGGGAACAGACACAGTCTTGTGTGCAATTTACC 300
QY	407 TCCGTCATATTTTATTTTCAATGTATATCAGTATTTCAATCTCTGGACGTAACTATC 466
Db	301 TCCGTCATATTTTATTTTCAATGTATATCAGTATTTCAATCTCTGGACGTAACTATC 360
QY	467 GATGCTACCGAAGACACACGAGCATTTTAAATCCAAACCCAAATCTCTTGGG 526
Db	361 GATGCTACCGAAGACACACGAGCATTTTAAATCCAAACCCAAATCTCTTGGG 420
QY	527 GCTAAGATTTCTCTGTTGTCTATCTGGGCAATTCATGTTCTTACTCTTTGGCTTAACATG 586
Db	421 GCTAAGATTTCTCTGTTGTCTATCTGGGCAATTCATGTTCTTACTCTTTGGCTTAACATG 480
QY	587 ATTCGACCAACAGCAGCCGAGAGACAAAGATGTGAAGAAATGCTCTTCTTAATCA 646
Db	481 ATTCGACCAACAGCAGCCGAGAGACAAAGATGTGAAGAAATGCTCTTCTTAATCA 540
QY	647 GAGTTCGGTCTAGTGTGGCATGAATATGTAATTTACATCTGTCAAGTCAATTTCTGGATT 706
Db	541 GAGTTCGGTCTAGTGTGGCATGAATATGTAATTTACATCTGTCAAGTCAATTTCTGGATT 600
QY	707 AATTTCTTAATTTGTTATTTGTATATACATCATTTTACAAAAGACGTACCGGTCATAC 766
Db	601 AATTTCTTAATTTGTTATTTGTATATACATCATTTTACAAAAGACGTACCGGTCATAC 660
QY	767 GTAAGACGAGGGGTGTAGTAAGTCCCAAGAAAAAGGTGAAGTCAAAATTTTCATT 826
Db	661 GTAAGACGAGGGGTGTAGTAAGTCCCAAGAAAAAGGTGAAGTCAAAATTTTCATT 720
QY	827 ATCATTTGCTGATTTCTTAATTTGTTGTCTTTCATTTTGGCCGAATTCCTTACCC 886
Db	

DB	721	ATGATTGCTGTAATTCCTTATTTGTTTGTTCCTTCCATTTTGCCCGAAATTCCTTACAC	780
QY	887	CTGAGCCAAACCCGGGATGTCCTTGACGTCGACCTGCTGAAAAATCTGCTTATGTGAAA	946
DB	781	CTGAGCCAAACCCGGGATGTCCTTGACGTCGACCTGCTGAAAAATCTGCTTATGTGAAA	840
QY	947	GAGACCACTCTGTGGTTAACTTCCCTTAATGATGATGCCCTGATCCGTTATCATTTTTC	1006
DB	841	GAGACCACTCTGTGGTTAACTTCCCTTAATGATGATGCCCTGATCCGTTATCATTTTTC	900
QY	1007	CTTTCAAGTCCTTGAGAAATCTCTTGATGATGATGACGAGTGGCCCAATTCCTCAACA	1066
DB	901	CTTTCAAGTCCTTGAGAAATCTCTTGATGATGATGACGAGTGGCCCAATTCCTCAACA	960
QY	1067	TCTCTGTCCAGACGACATAGGAAAAAGAACAGATGCTGTGTGACCCAAATGAGAGACT	1126
DB	961	TCTCTGTCCAGACGACATAGGAAAAAGAACAGATGCTGTGTGACCCAAATGAGAGACT	1020
QY	1127	CCAATGTAA 1135	
DB	1021	CCAATGTAA 1029	
RESULT 11			
AB052684			
LOCUS	AB052684	1029 bp	linear PRI 01-SEP-2001
DEFINITION	Homo sapiens mRNA for G1-coupled ADP receptor HOKR3, complete cds.		
ACCESSION	AB052684		
VERSION	AB052684.1	GI:14422409	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens		
REFERENCE	Homo sapiens		
AUTHORS	Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.		
TITLE	1		
JOURNAL	Takasaki, J., Kamohara, M., Saito, T., Matsumoto, M., Matsumoto, S.,		
MEDLINE	Ohashi, T., Soga, T., Matsushima, H. and Furuchi, K.		
AUTHORS	pharmacological comparison with another ADP receptor, the P2Y(1)		
TITLE	receptor		
JOURNAL	Mol. Pharmacol. 60 (3), 432-439 (2001)		
MEDLINE	2 (bases 1 to 1029)		
AUTHORS	Takasaki, J., Kamohara, M., Saito, T., Matsumoto, M., Matsumoto, S.,		
TITLE	Ohashi, T., Soga, T., Matsushima, H. and Furuchi, K.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (13-DEC-2000) Jun Takasaki, Yamanouchi pharmaceutical		
AUTHORS	Co., Ltd., Molecular Medicine Laboratories; 21 Miyukigaoka,		
TITLE	Tsukuba, Ibaraki 305-8585, Japan (E-mail: takasaki@yamanouchi.co.jp,		
JOURNAL	Tel:81-298-52-5111, Fax:81-298-52-5444)		
MEDLINE	Location/Qualifiers		
AUTHORS	1. 1029		
TITLE	/organism="Homo sapiens"		
JOURNAL	/db_xref="taxon:9606"		
MEDLINE	/chromosome="3"		
AUTHORS	/map="3q21-q25"		
TITLE	1. 1029		
JOURNAL	/gene="hork3"		
MEDLINE	1. 1029		
AUTHORS	/gene="hork3"		
TITLE	/codon_start=1		
JOURNAL	/product="G1-coupled ADP receptor HOKR3"		
MEDLINE	/protein_id="BAB60824.1"		
AUTHORS	/db_xref="GI:14422410"		
TITLE	/translation="MQAVDNTLSAPGNTSLCTRDYKTYOVLFPILYVLPFVGLITMG		
JOURNAL	LAMHIFQIRKSNFLLFLKATVTSIDMLITFEFKILSDKLTGRLPRTFVCVSTSV		
MEDLINE	ILFTWTYISISFLGILIDRQKTRTFKSNPNLIGAKILSVIAWFLSLPMN		
AUTHORS	ILYTRORDKKNVCKCSFLKSEFGVMEHIVYICQVLEIMFLIVCYLTITELR		
TITLE	SVYRSGVGVPRKRVKVFIIIAVFCIPVPHPARIPPTTSOTRDVPCATENTL		
JOURNAL	FVYESTLMLTSLNACDPLPIYFLCSPFRNLSLMLKPSNATSLISODNRKKQDDG		
MEDLINE	DPNEETW"		
BASE COUNT	285 a	234 c	186 g 324 t

Query Match 86.4%: Score 1027.4; DB 9; Length 1029;
 Best Local Similarity 99.9%: Pred. No. 7, 8e-242;
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 107 ATGCAACCGCTGCACAAATCTCACCCTGCGCTGGGAACACCACTGTGTGCACGAGAC 166
    |||||||
DB 1 ATGCACACCGCTGCACAACTCACCCTGCGCTGGGAACACCACTGTGTGCACGAGAC 60
QY 167 TACAAATCAACCCAGGCTCTTCCCACTGCTACACAGTGGCTTTTGTGTGGACTT 226
    |||||||
DB 61 TACAAATCAACCCAGGCTCTTCCCACTGCTACACAGTGGCTTTTGTGTGGACTT 120
QY 227 ATCAAAATGGCCCTGGGATGAGAGATTTCTTCAAAATCCGAGTAATCAACTTATT 286
    |||||||
DB 121 ATCAAAATGGCCCTGGGATGAGAGATTTCTTCAAAATCCGAGTAATCAACTTATT 180
QY 287 ATTTTCTTAAGAACACAGTCATTTCTGANTCTTCTCATGATTTGACTTTTCATCAA 346
    |||||||
DB 181 ATTTTCTTAAGAACACAGTCATTTCTGANTCTTCTCATGATTTGACTTTTCATCAA 240
QY 347 ATTCTTAGTGTGCAACACTGGGAGGAGGACGACACTGTTGTGTCAAGTTACC 406
    |||||||
DB 241 ATTCTTAGTGTGCAACACTGGGAGGAGGACGACACTGTTGTGTCAAGTTACC 300
QY 407 TCCGTCATATTTTATTTACAAATGTAATCATGATTTCTTCTGGGACTGATTAATC 466
    |||||||
DB 301 TCCGTCATATTTTATTTACAAATGTAATCATGATTTCTTCTGGGACTGATTAATC 360
QY 467 GATGCTACCAAGAACACGACGACGCAATTTAAACATCCAAACCCAAATCTCTGGG 526
    |||||||
DB 361 GATGCTACCAAGAACACGACGACGCAATTTAAACATCCAAACCCAAATCTCTGGG 420
QY 527 GCTAAGATTCCTCTGTTGTCATCTGGGCAATTCATGTTCTTCTTGGCTTAACATG 586
    |||||||
DB 421 GCTAAGATTCCTCTGTTGTCATCTGGGCAATTCATGTTCTTCTTGGCTTAACATG 480
QY 587 ATTCTGACCAACAGGACGCGAGAGACAGACAGTAATGTAAGAAATGCTTTCTTAAATCA 646
    |||||||
DB 481 ATTCTGACCAACAGGACGCGAGAGAGACAGTAATGTAAGAAATGCTTTCTTAAATCA 540
QY 647 GAGTTCGGCTAGTCTGTGGCAGTAATGTAATTAATCACTGTCTCAAGTCATTTTCTGANT 706
    |||||||
DB 541 GAGTTCGGCTAGTCTGTGGCAGTAATGTAATTAATCACTGTCTCAAGTCATTTTCTGANT 600
QY 707 AATTCTTAATTTGTTATGTAATGTAATCACTCATTAACAAAGAACTGACCGGCTACAC 766
    |||||||
DB 601 AATTCTTAATTTGTTATGTAATGTAATCACTCATTAACAAAGAACTGACCGGCTACAC 660
QY 767 GTAAGAACGAGGGGTGAGTAAAGTCCAGGAAAAAGGTGAAGCTCAAGTTTCATT 826
    |||||||
DB 661 GTAAGAACGAGGGGTGAGTAAAGTCCAGGAAAAAGGTGAAGCTCAAGTTTCATT 720
QY 827 ATCATTCCTGATTTCTTATTTGTTGTTCTTTCATTTTCCGCAATTTCTTACACC 886
    |||||||
DB 721 ATCATTCCTGATTTCTTATTTGTTGTTCTTTCATTTTCCGCAATTTCTTACACC 780
QY 887 CTGAGCCAAACCCGGAGTGTCTTGAAGTCAGTCTGAAATFACTGTTCTATGTGAAA 946
    |||||||
DB 781 CTGAGCCAAACCCGGAGTGTCTTGAAGTCAGTCTGAAATFACTGTTCTATGTGAAA 840
QY 947 GAGAGCACTGTGTTTAACTTCTTAAATGCAATGCTGATCCGTTCACTATTTTTC 1006
    |||||||
DB 841 GAGAGCACTGTGTTTAACTTCTTAAATGCAATGCTGATCCGTTCACTATTTTTC 900
QY 1007 CTTTGCAAGTCTTCACAAATTCCTTATAGTATGCTGAAGTCCCAATTTCTGCACA 1066
    |||||||
DB 901 CTTTGCAAGTCTTCACAAATTCCTTATAGTATGCTGAAGTCCCAATTTCTGCACA 960
QY 1067 TCTCTGTCCAGACATATAGGAAAAAAGAACAGAGATGCTGTGACCAATGAAGAGACT 1126
    |||||||
DB 961 TCTCTGTCCAGACATATAGGAAAAAAGAACAGAGATGCTGTGACCAATGAAGAGACT 1020
  
```

QY 1127 CCATGTAA 1135
 |||||||
 DB 1021 CCATGTAA 1029

RESULT 12
 AB083596 1029 bp DNA linear PRI 24-MAY-2002
 LOCUS
 DEFINITION Homo sapiens GPCR gene for putative G-protein coupled receptor,
 complete CDS, clone:hgPCR14.
 ACCESSION AB083596
 VERSION AB083596.1 GI:20152255
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens DNA, clone:hgPCR14.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Takeda,S., Kadowaki,S., Haga,T., Takeasu,H. and Mitaku,S.
 TITLE Identification of G protein-coupled receptor genes from the human genome sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1029)
 AUTHORS Takeda,S., Kadowaki,S., Haga,T., Takeasu,H. and Mitaku,S.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-2002) Shigeaki Takeda, Gunma University,
 Department of Biological and Chemical Engineering, Faculty of
 Engineering, 1-5-1, Kiryu, Gunma 376-8515, Japan
 (E-mail:stakeda@bce.gunma-u.ac.jp, Tel:+81-277-30-1434,
 Fax:+81-277-30-1434)

FEATURES
 source
 1..1029
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="hgPCR14"
 1..1029
 /gene="GPCR"
 1..1029
 /gene="GPCR"
 /note="predicted with SOSUI analysis"

CDS
 1/codon_start=1
 /product="putative G-protein coupled receptor"
 /protein_id="BAB89309.1"
 /db_xref="GI:20152255"
 /translation="MQAVDMLTSAPGNTSLCTPDYKITQVLPFLITVLPFVGLING
 LAMRIPOISKSNFIFLKNVYISDLMLITTPRILTSARKEPLRIFVQVTSV
 IFFMTISIFLGLTIDRYQNTTPREKTSNKNLGLAKISLVITAEFLLSLPMM
 ILVNRQPRDKNVKCSFLKSEFLVWHEIVNYICQVIFWNLIVICVYLITRELYR
 SYVTRGVGVPRKKNVKKVFIIVAFELFVPHFARIPYLSQTRDVEDCTAENTL
 FVKESTLMTISLNACLDPRIFYFLCKSFNSLSLSMLKCNASATLSQDRKKEDQGG
 DPHEETPM"

BASE COUNT 285 a 234 c 185 g 325 t
 ORIGIN

Query Match 86.4%: Score 1025.8; DB 9; Length 1029;
 Best Local Similarity 99.8%: Pred. No. 1, 9e-241;
 Matches 1027; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 107 ATGCAACCGCTGCACAAATCTCACCCTGCGCTGGGAACACCACTGTGTGCACGAGAC 166
    |||||||
DB 1 ATGCACACCGCTGCACAACTCACCCTGCGCTGGGAACACCACTGTGTGCACGAGAC 60
QY 167 TACAAATCAACCCAGGCTCTTCCCACTGCTACACAGTGGCTTTTGTGTGGACTT 226
    |||||||
DB 61 TACAAATCAACCCAGGCTCTTCCCACTGCTACACAGTGGCTTTTGTGTGGACTT 120
QY 227 ATCAAAATGGCCCTGGGATGAGAGATTTCTTCAAAATCCGAGTAATCAACTTATT 286
    |||||||
DB 121 ATCAAAATGGCCCTGGGATGAGAGATTTCTTCAAAATCCGAGTAATCAACTTATT 180
QY 287 ATTTTCTTAAGAACACAGTCATTTCTGANTCTTCTCATGATTTGACTTTTCATCAA 346
    |||||||
DB 181 ATTTTCTTAAGAACACAGTCATTTCTGANTCTTCTCATGATTTGACTTTTCATCAA 240
  
```

QY 347 ATTCTTAGTATGCCAACTGGGACAGACCACTGAACTTTGTGTCAAGTTACC 406
 |||||||
 Db 241 ATTCTTAGTATGCCAACTGGGACAGACCACTGAACTTTGTGTCAAGTTACC 300
 QY 407 TCGGTATATTTATTTACAAATGTAATGATATTTCAATCTCGGGACGATACATAC 466
 |||||||
 Db 301 TCGGTATATTTATTTACAAATGTAATGATATTTCAATCTCGGGACGATACATAC 360
 QY 467 GATCGGTACCAAGAACACACAGCCATTTTAAACATCCAAACCCCAAAATCTCTGGGG 526
 |||||||
 Db 361 GATCGGTACCAAGAACACACAGCCATTTTAAACATCCAAACCCCAAAATCTCTGGGG 420
 QY 527 GCTAAGATTCTCTGTGTGATCTGGGCAATGATGTTCTTACTCTTGTGCTTACATG 586
 |||||||
 Db 421 GCTAAGATTCTCTGTGTGATCTGGGCAATGATGTTCTTACTCTTGTGCTTACATG 480
 QY 587 ATTCTGCACCAAGACAGCCGAGAGACAAAGATGTGAAGAAATGCTCTTCTTAAATCA 646
 |||||||
 Db 481 ATTCTGCACCAAGACAGCCGAGAGACAAAGATGTGAAGAAATGCTCTTCTTAAATCA 540
 QY 647 GAGTGGGTCTAGTCTGGCAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 706
 |||||||
 Db 541 GAGTGGGTCTAGTCTGGCAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 707 AATTTCTTAATTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 766
 |||||||
 Db 601 AATTTCTTAATTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 767 GTAGAAGAGAGGGGTAGTGAAGTCCCGAGGAAAAAGGTGAAGGTAAGTTTCATT 826
 |||||||
 Db 661 GTAGAAGAGAGGGGTAGTGAAGTCCCGAGGAAAAAGGTGAAGGTAAGTTTCATT 720
 QY 827 ATCATGCTGATCTTATTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 886
 |||||||
 Db 721 ATCATGCTGATCTTATTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
 QY 887 CTAGAGCAAAACCCGGATGCTTTGACTGACGCTGTAAGTAATCTCTGTTCTATGTA 946
 |||||||
 Db 781 CTAGAGCAAAACCCGGATGCTTTGACTGACGCTGTAAGTAATCTCTGTTCTATGTA 840
 QY 947 GAGAGCACTGTGTTGTTAACTCTTAATGATGATGATGATGATGATGATGATGATGAT 1006
 |||||||
 Db 841 GAGAGCACTGTGTTGTTAACTCTTAATGATGATGATGATGATGATGATGATGATGAT 900
 QY 1007 CTTTGCAGAGTCTTCAAGAAATCTTGATAGTATGCTGAAGTGCCTCAATTTCTGACA 1066
 |||||||
 Db 901 CTTTGCAGAGTCTTCAAGAAATCTTGATAGTATGCTGAAGTGCCTCAATTTCTGACA 960
 QY 1067 TCTCTGCCAGGACAAATGAGAAAAAGAACAGAGATGTTGTAACCAATGAAAGACT 1126
 |||||||
 Db 961 TCTCTGCCAGGACAAATGAGAAAAAGAACAGAGATGTTGTAACCAATGAAAGACT 1020
 QY 1127 CCAATGTAA 1135
 |||||||
 Db 1021 CCAATGTAA 1029
 RESULT 13
 BC027381 2151 bp mRNA 1linear ROD 07-AUG-2002
 LOCUS BC027381
 DEFINITION Mus musculus, putative G-protein coupled 12, clone
 ACCESSION MGC:36953 IMAGE:4947057, mRNA, complete cds.
 VERSION BC027381
 KEYWORDS house mouse.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2151)
 AUTHORS Strausberg, R.
 TITLE Direct Submission

JOURNAL

Submitted (04-Apr-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail@nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyik, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 60 Row: d Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers

source

1..2151
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:36953 IMAGE:4947057"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."
 /clone_1lb="NCI_CGAP_Mame"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 217..1260
 /codon_start=1
 /product="putative G-protein coupled receptor p2y, G-protein coupled 12"
 /protein_id="AAH27381.1"
 /db_xref="GI:20073303"
 /db_xref="GI:20073303"
 /db_xref="locusID:70839"
 /translation="MDVGVNTSNTAFSPSTSLCYKDRKITGVLEPLLYTFFA
 GLITSLARIFQJLRKSNFIFLKNVVISDLMLTFPEKILSDALGAPLTLV
 CQVTSVTFYFMWYISIFLGLITIDRYLKTFRPTSPSNNLGAFLISVYAWMFL
 ISLPMWILNRRPKDKDYKCSFLSEGLWHELVNLCQVTFMILNLIYVCSLI
 TKELRSYVRFGSAKVPKKNVAVFIIIAVFIICPVPFHRARIPYLSORAVEDC
 SAENLTFYKESTLWLSINCLDPFIIFLFLCKSRNLSLRSNSSTSGTKKK
 GQGEPSSEPRFM"
 BASE COUNT 717 a 456 c 360 g 618 t
 ORIGIN

CDS

Query Match 61.7%; Score 732.2; DB 10; Length 2151;
 Best local Similarity 82.2%; Pred. No. 2,7e-169;
 Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;
 QY 114 CGGTGACAAATCTCACTCTGGCGCTGGGAACACAGCTGTGCACCAAGACTACAAA 173
 || |||||
 Db 242 CCTGACCAATACCACTTCTCCCTGGGACACGACCACTGTCGACAGACTACAGA 301
 QY 174 TCACCCAGAGTCTCTCCACAGCTCTACAGTCTCTGTTTGTGGACTATACAA 233
 |||||||
 Db 302 TCACCCAGAGTCTCTCCACAGCTCTACAGTCTCTGTTTGTGGACTATACAGA 361
 QY 234 ATGCGCTGGGAGATGAGATTTCTTCAATCCGAGTAATCAACTTATATTTTC 293
 || |||||
 Db 362 ACAGCTTGGCAATGAGGATTTCTTCAAGTCCGAGTAATCAACTTATATTTTC 421
 QY 294 TTAGAACAACAGTCTTGTGATCTGATGATGATGATGATGATGATGATGATGATGAT 353
 |||||||
 Db 422 TTAGAACAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
 QY 354 GTGATGCCAACTGGGACAGACCACTGTAAGTCTGATGATGATGATGATGATGATGAT 413

```

Db 482 GTGATGCTAAACCTGGAGCCGGCCCTGAGAACCTGGTGTGCAATGTTCTTCAGCA 541
Qy 414 TATTTATTTTACATGATATATCATGATTTTCATTCCTGGAGATATACATATCATGCT 473
Db 542 CATTTTATTTTACATGATATATCATGATTTTCATTCCTGGAGATATACATATCATGCT 601
Qy 474 ACCAAGACCAACAGGACATTTTAAACATCCAAACCCCAAAATCTCTTGGGGCTAAGA 533
Db 602 ACCGAGACCAACAGGACATTTTAAACATCCAAACCCCAAAATCTCTTGGGGCTAAGA 661
Qy 534 TTTCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593
Db 662 TTTCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
Qy 594 CCAACAGGACCAACAGGACATTTTAAACATCCAAACCCCAAAATCTCTTGGGGCTAAGA 653
Db 722 CCAACAGGACCAACAGGACATTTTAAACATCCAAACCCCAAAATCTCTTGGGGCTAAGA 781
Qy 654 GTCTAGTCTGGCATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
Db 782 GTCTAGTCTGGCATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
Qy 714 TATTTATTTTACATGATATATCATGATTTTCATTCCTGGAGATATACATATCATGCT 773
Db 842 TATTTATTTTACATGATATATCATGATTTTCATTCCTGGAGATATACATATCATGCT 901
Qy 774 CGAGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 833
Db 902 CAAGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
Qy 834 CTGATATCTTATTTATTTGTTTCTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 893
Db 962 CTGATATCTTATTTATTTGTTTCTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 1021
Qy 894 AAACCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953
Db 1022 AAACCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
Qy 954 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013
Db 1082 CCGATATCTTATTTATTTGTTTCTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 1141
Qy 1014 AGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073
Db 1142 AGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
Qy 1074 CCGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
Db 1199 CTGGGACCAACAGGACATTTTAAACATCCAAACCCCAAAATCTCTTGGGGCTAAGA 1258
Qy 1134 AAACAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1166
Db 1259 AGA-ACATTAACCAAGGGGTGCTTCATCTTTT 1290

RESULT 14
BC025428 2422 bp mRNA linear ROD 07-NDG-2002
LOCUS Mus musculus, clone MGC:29342 IMAGE:5035481, mRNA, complete cds.
ACCESSION BC025428
VERSION BC025428.1 GI:19343840
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scleroglossa; Muridae; Mus.
REFERENCE
1 (bases 1 to 2422)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

```

REMARK
COMMENT
USA
NHI-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhi.nih.gov
Arter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Grantham, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maki, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Staudt, P., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov
Series: IRAK Plate: 39 Row: 1 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES
source
location/Qualifiers
1..2422
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C2ECH 11"
/clone="MGC:29342 IMAGE:5035481"
/tissue="type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MTV."
/clone_id="NCI_CGAP_1429"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
422..1465
/codon_start=1
/product="Unknown (protein for MGC:29342)"
/protein_id="AAH25428.1"
/db_xref="GI:19343841"
/translaction="MDMPGVNTTSANTFSPGFTLCVRDYKIRIOVLEPLTYLFFA
GLTNSLAMIIFQIRKSNFIPLKTVISLIMLITLPPKISLDAKLAGPLRTY
ISUPNMLNRRPKDQVYKCSFLKSEGLVHWIVATQVLEWIFLIVCYSLI
TKELYRSYVTRGSARKVKKVNVKVFIIIAVEFICFPHFARIPTYLQTRAVFDC
SAENTLIFYVESTIMLSLNAICLDPFIYFLFKSFRNLSMLSCNSTSGTKNKK
GOEGESEPSETPM"

BASE COUNT 781 a 523 c 419 g 699 t
ORIGIN
Query Match 61.7%; Score 732.2; DB 10; Length 2422;
Best Local Similarity 82.2%; Pred. No. 2,76-169;
Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;
Qy 114 CCGTGGAAATCTACCCCTGCGGCTGGGAACACCAAGCTGTGACACGACATCAAA 173
Db 447 CCGTGGAAATCTACCCCTGCGGCTGGGAACACCAAGCTGTGACACGACATCAAA 506
Qy 174 TCACCCAGGTCCTCTCCACAGTCTCTACAGTCTCTCTTTTGTGAGCATTAACAA 233
Db 507 TCACCCAGGTCCTCTCCACAGTCTCTACAGTCTCTCTTTTGTGAGCATTAACAA 566
Qy 234 ATGGCTGGCATGAGATTTTCTTCAATTCGGAGTAAATCAAACTTATATTTTC 293
Db 567 ACAGCTTGGCATGAGATTTTCTTCAATTCGGAGTAAATCAAACTTATATTTTC 626
Qy 294 TTAGAACACGATCTTCTGATCTTCATGATCTGATCTGATCTTCCATCAAAATCTTA 353
Db 627 TTAGAACACGATCTTCTGATCTTCATGATCTGATCTTCCATCAAAATCTTA 686

```

QY	354	GTGATGCGCAAACTGGGAGACAGCACACTAGAACTTTTGTGTCTCAAGTTACCTCCGTCA	413
Db	687	GGATGCTAAACTGGGGAGCCGGGCGCTCTGGAACCTTGGGTGCCAAGTTACTTCAGTCA	746
QY	414	TATTTTATTTTCCAAATGATATATCAGTATTTTCAATCTCTGGGACTGATATATCGATCGCT	473
Db	747	CATTTTATTTTCCAAATGATATATCAGTATATCGTTCTCTGGGTTGTAATACCATTTGACCGCT	806
QY	474	ACCAGAAAGACACACAGGCCATTTTAAACAATCCCAACCCAAAATCTCTTGGGGGCTAA	533
Db	807	ACTGGAAGACACACAGGCCATTTTAAACAATCCCAACCCAAAATCTCTTGGGGGCTAA	866
QY	534	TTCTCTCTCTTGTGCATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATGATCTGGA	593
Db	867	TTCTCTCTCTTGTGCATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATGATCTGGA	926
QY	594	CCAAACGAGGAGCGGAGAGACAGAAATGTGAGAAATAGCTTTCCCTTAAATCGAGCTTGG	653
Db	927	CCAAACGAGGAGCGGAGAGATTAAGAGAGCTTAACAAATGTTCTTTTAAAGTCAGACTTGG	986
QY	654	GTCCTACTGGGAGTGAATAGTAAATTTACATCTGTCAAGTCAATTTTCTGGATTAATTTCT	713
Db	987	GTCCTACTGGGAGTGAATAGTAAATTTACATCTGTCAAGTCAATTTTCTGGATTAATTTCT	1046
QY	714	TATTTGTTATTTGTATGTTATATACATCTATTTACAAGAAACTGTACCGGCTATACGTAAAGA	773
Db	1047	TATTTGTTATTTGTATGTTATATACATCTATTTACAAGAAACTGTACCGGCTATATGTACGA	1106
QY	774	CGAGGGGCTGTAGGTAAAGTCCGCCAGGAAAGAGGTGAAGCTCAAGATTTTCATATCATTTG	833
Db	1107	CAAGGGGCTGTAGGTAAAGTCCGCCAGGAAAGAGGTGAAGCTCAAGATTTTCATATCATTTG	1166
QY	834	CTGTATTTCTTTATTTGTTTGTTCCTTTCCATTTTCCCGGAATTCCTTACACCCGTAGCC	893
Db	1167	CTGTATTTCTTTATTTGTTTGTTCCTTTCCATTTTCCCGGAATTCCTTACACCCGTAGCC	1226
QY	894	AAACCCGGGATGCTCTTGTGACTGACAGTCCGAAATATCTGTGTTCTATGTGAAGAGAGA	953
Db	1227	AAACCCGGGCGCTCTTGTGACTGACAGTCCGAAATATCTGTGTTCTATGTGAAGAGAGA	1286
QY	954	CTCTGTGTTAACTTCCCTTAAATGATGCGCTGGAATCCGTTCAATCTATTTTCTTTGCA	1013
Db	1287	CCCTATTTGCTTACGTCACAGAACCGCTGCTTGATTCATTCATCTTTTCTTTTGTGCA	1346
QY	1014	ACTCTTCAGAAATTCCTTGATTAAGTATGCTGAAGTGGCCCCCAATTCGTGCAATCTCTGT	1073
Db	1347	ACTCTTCAGAAATTCCTTGATTAAGTATGCTGAAGTGGCCCCCAATTCGTGCAATCTCTGT	1403
QY	1074	CCGAGGACAAATGTAGAAAAAACAAGAGATGTGTGACCCCAATGAGAAGAGATCTCAATGT	1133
Db	1404	CTGGGACAAACAAGAAAGAAAGAGACAAAGAGATGTGTGACCCCAATGAGAAGAGATCTCAATGT	1463
QY	1134	AAACCAATTTTAACTAAGGAAATATTTTCAATCTCT	1166
Db	1464	AGA-ACATTTACCCAAAGGGCTGCTTCAAGTCTTT	1495
RESULT 15			
AC122038			
LOCUS	AC122038	211964 bp	DNA linear HTG 27-MAY-2002
DEFINITION	Mus musculus chromosome UNK clone RP24-427G24, WORKING DRAFT		
ACCESSION	AC122038		
VERSION	AC122038.1	GI:21040290	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	McPherson, J.D. and Waterston, R.H.		
TITLE	The sequence of Mus musculus clone		

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 211964)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
(bases 1 to 211964)
McPherson,J.D. and Waterston,R.H.
REFERENCE Direct Submission
AUTHORS Submitted (27-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
JOURNAL

COMMENT ----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project information
-----
Center project name: M_BB0427G24
-----
Summary Statistics
-----
Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-primer EN; 0% of reads
Chemistry: Dye-terminator Big_Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210721 bases at least Q40
Consensus quality: 211170 bases at least Q30
Consensus quality: 211464 bases at least Q20
Insert size: 195000; agarose-fp
Insert size: 211764; sum-of-contigs
Quality coverage: 12.85 in Q20 bases; agarose-fp
Quality coverage: 10.63 in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently
consists of 3 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*      *      *
*      1      1402: contig of 1402 bp in length
*      *      1502: gap of unknown length
*      *      1503          91430: contig of 89928 bp in length
*      *      91431          91530: gap of unknown length
*      *      91531          211964: contig of 120434 bp in length.
*      *      Location/Qualifiers
*      *      1. .211964
*      *      /organism="Mus musculus"
*      *      /db_xref="taxon:10090"
*      *      /chromosome="YUNK"
*      *      /clone="RP24-427G24"
*      *      1. .1402
*      *      /note="assembly_name:Contig5"
*      *      1503. .91430
*      *      /note="assembly_name:Contig6"
*      *      91531. .211964
*      *      /note="assembly_name:Contig7"
BASE COUNT   62887 a 43533 c 43160 g 62052 t       212 others
ORIGIN
Query Match           61.7%; Score 732.2; DB 2; Length 211964;
Best Local Similarity 82.2%; Pred. No.2.5e+16;
Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;

OY    174 TCACCCAGGCTCCTTCACACGTCTACACGTCTGTTTGGACTTAACAA 233
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 108731 CCTGCACAATACCACTTCTCCCTGGGACCAGCACCTGTGGCTCAGAGCTACAAGA 108790
        TACCCAGGCTCCTTCACACGTCTACACGTCTGTTTGGACTTAACAA 108850
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| |

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 22:50:25 ; Search time 322 Seconds
(without alignments)
8301.621 Million cell updates/sec

Title:	US-09-835-922-1
Perfect score:	1187
Sequence:	1 ggtaccgcgagatcctctag.....tgygtcgaactcgttaaa 1187

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1187	100.0	1187	24	ABA92641	Human SP168 receptor
2	1160	97.7	1893	22	AAK52382	Human polynucleotide
3	1156.8	97.5	2249	20	AAV69760	EBV-induced G-protein
4	1156.8	97.5	2286	21	AAA27126	Human 15625 receptor
5	1153.4	97.2	1332	22	AAAS9679	Human DNA for protein
6	1151.4	97.0	1330	22	AAAS08695	Human DNA for protein
7	1137.4	95.8	1328	22	AAD08700	Human p2-purification
8	1108.8	93.4	2272	21	AA271127	Macaque ortholog of
9	1093.8	92.1	1101	24	AA168803	Human PAWR-3 DNA f

15	252.4	21.3	2416	24	ABN6152
44	253	21.3	1011	22	ABF57655
45	263.2	22.2	356	24	ABA8542
42	279	23.5	302	20	AAK51843
41	294.8	24.8	1591	24	ABL90619
40	295.4	24.9	1474	21	AAZ60118
39	295.4	24.9	1474	21	AAZ60118
38	295.4	24.9	1474	17	AAI33904
37	302.4	25.5	1488	19	AAI72559
36	304	25.6	1488	19	AAV59144
35	305.6	25.7	1124	23	AAAF6912
34	305.6	25.7	1124	22	AAAF6912
33	305.6	25.7	1124	20	AAK53585
32	305.6	25.7	1002	21	AAA46027
31	305.6	25.7	1002	21	AAA46027
30	307.2	25.9	2760	24	AAK58051
29	307.2	25.9	1812	21	AAA33616
28	307.2	25.9	1484	22	AAAF87601
27	307.2	25.9	1484	22	AAAF87601
26	307.2	25.9	1484	22	AAAF87601
25	307.2	25.9	1484	22	AAAF87601
24	307.2	25.9	1484	22	AAAF87601
23	307.2	25.9	1484	22	AAAF87601
22	307.2	25.9	1484	22	AAAF87601
21	307.2	25.9	1484	22	AAAF87601
20	307.2	25.9	1484	22	AAAF87601
19	307.2	25.9	1484	22	AAAF87601
18	307.2	25.9	1484	22	AAAF87601
17	307.2	25.9	1484	22	AAAF87601
16	307.2	25.9	1484	22	AAAF87601
15	307.2	25.9	1484	22	AAAF87601
14	307.2	25.9	1484	22	AAAF87601
13	307.2	25.9	1484	22	AAAF87601
12	307.2	25.9	1484	22	AAAF87601
11	307.2	25.9	1484	22	AAAF87601
10	307.2	25.9	1484	22	AAAF87601
9	307.2	25.9	1484	22	AAAF87601
8	307.2	25.9	1484	22	AAAF87601
7	307.2	25.9	1484	22	AAAF87601
6	307.2	25.9	1484	22	AAAF87601
5	307.2	25.9	1484	22	AAAF87601
4	307.2	25.9	1484	22	AAAF87601
3	307.2	25.9	1484	22	AAAF87601
2	307.2	25.9	1484	22	AAAF87601
1	307.2	25.9	1484	22	AAAF87601

ALIGNMENTS

RESULT 1	ID	AB092641	standard; cDNA; 1187 BP.
XX	AC	AB092641;	
XX	DT	25-MAR-2002 (first entry)	
XX	DE	Human SP168 receptor encoding cDNA SEQ ID NO:1.	
XX	KW	Human: SP168 receptor; mammalian G-protein coupled receptor; GPCR;	
XX	KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;	
XX	KW	Huntington's disease; amyotrophic lateral sclerosis; ALS; MS; receptor	
XX	KW	multiple sclerosis; gene; ss.	
XX	OS	Homo sapiens.	
XX	FT	Location/Qualifiers	
XX	FT	107..1135	
XX	FT	/*tag= a	
XX	FT	/product= "SP168 receptor"	
XX	PN	US2001046497-A1.	
XX	PD	29-NOV-2001.	
XX	PF	16-APR-2001; 2001US-0835922.	
XX	PR	21-APR-2000; 2000US-199041P.	
XX	PA	(ZHANG) ZHANG F L.	
XX	PA	(LUOL) LUO L.	

PA (GUST/) GUSTAFSON E.
 PA (LIU/) LIU Y.
 PA (CHEN/) CHEN G.
 PI Zhang FL, Luo L, Gustafson E, Liu Y, Chen G;
 DR WPI: 2002-082414/11.
 DR P-PSDB; AB805031.
 XX
 PT Identifying modulators of mammalian G-Protein Coupled Receptor SP168,
 PT useful for treating Parkinson's Disease, Alzheimer's Disease,
 PT Huntington's Disease, amyotrophic lateral sclerosis and multiple
 PT sclerosis -
 XX
 PS Example 1; Page 13-14; 16pp; English.

CC The present invention describes a method for identifying agonists and
 CC antagonists of mammalian G-Protein Coupled Receptor (GPCR) SP168 which
 CC may be used to treat neurodegenerative disorders. The method comprises:
 CC (a) contacting a mammalian SP168 receptor (or a functional fragment) in
 CC the presence of a known amount of a labeled SP168 receptor ligand with
 CC a sample to be tested for the presence of the SP168 receptor ligand with
 CC antagonist; and (b) measuring the amount of labeled SP168 ligand
 CC specifically bound to the receptor (the SP168 receptor agonist or
 CC antagonist in the sample is identified by measuring the difference in
 CC binding of the labeled SP168 receptor ligand to the receptor, compared
 CC to what would be measured in the absence of such agonist or antagonist).
 CC The method is used to detect agonists and antagonists (especially
 CC antibodies) of the SP168 GPCR which may be used to treat a
 CC neurodegenerative disorder, such as Parkinson's Disease, Alzheimer's
 CC Disease, Huntington's Disease, amyotrophic lateral sclerosis (ALS) and
 CC multiple sclerosis (MS) in mammals, especially humans. The present
 CC sequence encodes the human SP168 receptor which can be used in the
 CC method of the invention.

SQ Sequence 1187 BP; 340 A; 270 C; 212 G; 365 T; 0 other;

Query Match 100.0%; Score 1187; DB 24; Length 1187;
 Best Local Similarity 100.0%; Pred. No. 1e-304;
 Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCCGGGATCTCTAGAGTCGACCTGACGCTCAATTAATCTACTACTGATA 60
 DB 1 GGTACCCGGGATCTCTAGAGTCGACCTGACGCTCAATTAATCTACTACTGATA 60
 QY 61 CATTCAAACCTCCAGAAATACAGTATACAGTAACCAACCAAGCCGCTGA 120
 DB 61 CATTCAAACCTCCAGAAATACAGTATACAGTAACCAACCAAGCCGCTGA 120
 QY 121 CAATCTACCTCTGCGCTGGAAACACAGTCTGTGCACGACAGATACAAATACACCA 180
 DB 121 CAATCTACCTCTGCGCTGGAAACACAGTCTGTGCACGACAGATACAAATACACCA 180
 QY 181 GGTCTCTTCCACACGTGCTACACGTCTGTGTTTGTGACCTTATCAAAATGGCT 240
 DB 181 GGTCTCTTCCACACGTGCTACACGTCTGTGTTTGTGACCTTATCAAAATGGCT 240
 QY 241 GGGATGAGGATTTCTTCAATCCGAGATACAAATCAATTTATTTTCTTAAGAA 300
 DB 241 GGGATGAGGATTTCTTCAATCCGAGATACAAATCAATTTATTTTCTTAAGAA 300
 QY 301 CACAGTATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAAATTTCTAGTAGTC 360
 DB 301 CACAGTATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAAATTTCTAGTAGTC 360
 QY 361 CAAACTGGGAACAGACACACTGAGAACTTTGTGTCAAGTATCTCCGTCATATTTTA 420
 DB 361 CAAACTGGGAACAGACACACTGAGAACTTTGTGTCAAGTATCTCCGTCATATTTTA 420
 QY 421 TTTCACAATGTATATAGTATTTTCATTTCTGAGCATGATATGATGATGATGATGATG 480
 DB 421 TTTCACAATGTATATAGTATTTTCATTTCTGAGCATGATATGATGATGATGATGATG 480

QY 481 GACCACGAGCATTTAAACATCCACCCCAAAATCTCTGTGGGCGCTAAGATTCCTC 540
 DB 481 GACCACGAGCATTTAAACATCCACCCCAAAATCTCTGTGGGCGCTAAGATTCCTC 540
 QY 541 TGTGTGATCATGGGATTCATGTTCTTACCTCTTTCGCTAACATGATTTGACCAACAG 600
 DB 541 TGTGTGATCATGGGATTCATGTTCTTACCTCTTTCGCTAACATGATTTGACCAACAG 600
 QY 601 GCACCCGAGAGACAAAGATGTGAAGAAATGCTCTTTCCTTAATCAGATGCTGCTAGT 660
 DB 601 GCACCCGAGAGACAAAGATGTGAAGAAATGCTCTTTCCTTAATCAGATGCTGCTAGT 660
 QY 661 CTGGCATGAATAGTAAATATACATCTGTCAAGTCATTTTGTGATTAATTTATTTGT 720
 DB 661 CTGGCATGAATAGTAAATATACATCTGTCAAGTCATTTTGTGATTAATTTATTTGT 720
 QY 721 TATGTATGTTTAACTCAATTAACAAAGACAGTACCGGTCAATGATGAACGAGGG 780
 DB 721 TATGTATGTTTAACTCAATTAACAAAGACAGTACCGGTCAATGATGAACGAGGG 780
 QY 781 TGTAGTAAAGTCCCGAGAAAAGGTGAACGTCAAGTTTCATTAATTCATTCGTAT 840
 DB 781 TGTAGTAAAGTCCCGAGAAAAGGTGAACGTCAAGTTTCATTAATTCATTCGTAT 840
 QY 841 CTTTATTTGTTTGTTCCTTCCATTTTGGCCGAATTCCTTACACCTGAGCCAAACCG 900
 DB 841 CTTTATTTGTTTGTTCCTTCCATTTTGGCCGAATTCCTTACACCTGAGCCAAACCG 900
 QY 901 GGATGCTTTGACAGCTGCTGTAATTAATCTGTTTATGTGAAGAGACAGCTGTG 960
 DB 901 GGATGCTTTGACAGCTGCTGTAATTAATCTGTTTATGTGAAGAGACAGCTGTG 960
 QY 961 GTTAATCTCTTAAAGCATAGCTGATCCGTTTCATTAATTTTCCCTTGAAGTCCT 1020
 DB 961 GTTAATCTCTTAAAGCATAGCTGATCCGTTTCATTAATTTTCCCTTGAAGTCCT 1020
 QY 1021 CAGAAATTCCTGTAGTATAGTATGCTGAAGTCCCAATTCGCAACATCTGTGCCAGGA 1080
 DB 1021 CAGAAATTCCTGTAGTATAGTATGCTGAAGTCCCAATTCGCAACATCTGTGCCAGGA 1080
 QY 1081 CAATAGGAAAAAAGACAGAGATGTGTGACCCCAATGAAGAGACATCAATGTAAACAA 1140
 DB 1081 CAATAGGAAAAAAGACAGAGATGTGTGACCCCAATGAAGAGACATCAATGTAAACAA 1140
 QY 1141 TTAACTAAGGAAATATTTCAATCTCTTGTGTTGAGAACTGTTTAA 1187
 DB 1141 TTAACTAAGGAAATATTTCAATCTCTTGTGTTGAGAACTGTTTAA 1187

RESULT 2
 ID AAK52382 standard; cDNA; 1893 BP.
 XX AAK52382;
 AC
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide seq. ID NO 927.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 XX
 PN MO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 PR
 XX (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Xue Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue Q, Yang Y, Wehrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM79249.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS
 PS Claim 1; Page 3051-3052; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoietic regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX
 SQ Sequence 1893 BP; 620 A; 393 C; 318 G; 560 T; 2 other:
 Query Match 97.7%; Score 1160; DB 22; Length 1893;
 Best Local Similarity 100.0%; Pred. No. 1,8e-297;
 Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 CTGACGCTGCATATCTACTACTGATATCAAACTCCGCAATCAACAGTT 87
 DB 210 CTGCGGCTGCAATATCTACTACTGATATCAAACTCCGCAATCAACAGTT 269
 QY 88 ATCAGGTACCAACAAAGAAATGCAAGCGGTGACAACTCTGCGCTGGGGAAC 147
 DB 270 ATCAGGTACCAACAAAGAAATGCAAGCGGTGACAACTCTGCGCTGGGGAAC 329
 QY 148 CAGTCTGTGACACAGAGCTACAAATACCCAGGTCCTCCACAGCTCTACACTGT 207
 DB 330 CAGTCTGTGACACAGAGCTACAAATACCCAGGTCCTCCACAGCTCTACACTGT 389
 QY 208 CCGTGTCTGTGACTTATCACAATGCGCTGGCATGAGAGATTTCTTCAATCG 267
 DB 390 CCGTGTCTGTGACTTATCACAATGCGCTGGCATGAGAGATTTCTTCAATCG 449
 QY 268 GAGTAAATCAACATTTATTTCTTAAGAACACAGCTCTTGTGATCTTCATAT 327
 DB 450 GAGTAAATCAACATTTATTTCTTAAGAACACAGCTCTTGTGATCTTCATAT 509
 QY 328 TCTGACTTTTCCATCAAAATTTCTAGTATGCCAACTGGGAAACAGACACTGAGA 387
 DB 510 TCTGACTTTTCCATCAAAATTTCTAGTATGCCAACTGGGAAACAGACACTGAGA 569
 QY 368 TTTTGTGTCAAGTATCTCGCTATATTTATTTCAATGATGATGATGATTTCA 447
 DB 570 TTTTGTGTCAAGTATCTCGCTATATTTATTTCAATGATGATGATGATTTCA 629
 QY 448 CCGGAGCTGATATCTATCTGATCTGACAGAGACCAAGGCAATTAACATCA 507
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 630 CTTGGAGCTGATATCTATCTGATCTGACAGAGAACACACAGCCATTTAAACATCA 689
 QY 508 CCCCCAAAATCTCTGGGGGGCTAGATTTCTCTGTGTGATCTGGGCAATTCATGTTCT 567
 DB 630 CCCCCAAAATCTCTGGGGGGCTAGATTTCTCTGTGTGATCTGGGCAATTCATGTTCT 749
 QY 568 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACGCGAGACAAAGATGTGAAGA 627
 DB 750 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACGCGAGACAAAGATGTGAAGA 809
 QY 628 ATGCTCTTCTCTTAATATGAGATTTCTGCTGATCTGCGATGAAATGATTAATTCAT 687
 DB 810 ATGCTCTTCTCTTAATATGAGATTTCTGCTGATCTGCGATGAAATGATTAATTCAT 869
 QY 688 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGTTATGATCACTATCA 747
 DB 870 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGTTATGATCACTATCA 929
 QY 748 AGAAGTACCGGCTATACAGTAAGAGAGGAGGCTAGGTAAGTCCCGAGAAAAAGT 807
 DB 930 AGAAGTACCGGCTATACAGTAAGAGAGGAGGCTAGGTAAGTCCCGAGAAAAAGT 989
 QY 808 GAGCTCAAAATTTCTATATCATTTCTGATCTGATCTTTATTTGTTCTTCTTCATTT 867
 DB 990 GAGCTCAAAATTTCTATATCATTTCTGATCTGATCTTTATTTGTTCTTCTTCATTT 1049
 QY 868 TGCCCGAATTCCTTACACCCGAGCCAAACCGGAGTGTGATGATGATGATGATGATA 927
 DB 1050 TGCCCGAATTCCTTACACCCGAGCCAAACCGGAGTGTGATGATGATGATGATA 1109
 QY 928 TACTGTCTTATGTAAGAGAGACATCTGTGTTAACTTCTTAATGATGATGATGATA 987
 DB 1110 TACTGTCTTATGTAAGAGAGACATCTGTGTTAACTTCTTAATGATGATGATA 1169
 QY 988 TCCGTTCAATTTCTTCTTCTTCTTCTGAAATCTTCTGAAATCTTCTGAAATCTTCT 1047
 DB 1170 TCCGTTCAATTTCTTCTTCTTCTTCTGAAATCTTCTGAAATCTTCTGAAATCTTCT 1229
 QY 1048 GTGCCCAATTCCTGCAATCTCTGCTCCAGAGCAATAGGAAAGAAACAGAGATGAG 1107
 DB 1230 GTGCCCAATTCCTGCAATCTCTGCTCCAGAGCAATAGGAAAGAAACAGAGATGAG 1289
 QY 1108 TGACCCAAATGAGAGACTCCATGTAAACAAATTAAGTAAGAAATATTTCAATCTCT 1167
 DB 1290 TGACCCAAATGAGAGACTCCATGTAAACAAATTAAGTAAGAAATATTTCAATCTCT 1349
 QY 1168 TGTGTTCAAGACTCGTTAAA 1187
 DB 1350 TGTGTTCAAGACTCGTTAAA 1369
 RESULT 3
 AAV69760
 ID AAV69760 standard; cDNA; 2249 BP.
 XX AAV69760;
 XX 22-FEB-1999 (first entry)
 DE EBV-induced G-protein coupled receptor (EBI-2) encoding cDNA.
 XX
 XX EBV-induced G-protein coupled receptor (EBI-2) encoding cDNA.
 KW EBV-induced G-protein coupled receptor; EBI-2; Epstein-Barr Virus; ulcer;
 KW endothelium-differentiation gene; EDS-1-like G-protein coupled receptor;
 KW recombinant; agonist; asthma; Parkinson's disease; heart failure; asthma;
 KW hypotension; urinary retention; osteoporosis; antagonist; hypertension;
 KW angina pectoris; myocardial infarction; allergy; psychosis; depression;
 KW migraine; vomiting; stroke; eating disorder;migraine headache; cancer;
 KW prostatic hypertrophy; detection; drug screening; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 226..1255

XX OS Homo sapiens. Location/Qualifiers
 XX FH Key 261.1289
 XX FT CDS /*tag-
 FT /transl_except= (pos:297..299, aa:Asn)
 FT /product= "15625 receptor protein"
 XX PN MO200028028-A1.
 XX PD 18-MAR-2000.
 XX PF 05-NOV-1999; 99WO-US25956.
 XX PR 06-NOV-1998; 98US-0187134.
 XX PR 25-AUG-1999; 99US-0382918.
 XX PA (MILL-) MILENNIUM PHARM INC.
 XX PI Glucksmann MA, Gu W, Welch NS;
 XX DR WPI: 2000-376543/32.
 XX DR P-PSDB: AAY94444.
 XX PT Identifying an agent modulating the level or activity of G-protein
 PT coupled receptor useful for screening a cell derived from a subject
 PT having disorders such as anaemia, neutropenia and thrombocytopenia -
 XX PS Disclosure; Page 89-90; 97pp; English.
 XX CC The 15625 receptor protein is a novel G-coupled protein receptor (GPCR).
 CC The present sequence is the cDNA for this protein and was isolated by
 CC screening a human cDNA library with sequences homologous to other GPCRs.
 CC The 15625 receptor protein is expressed in the glial cells of the brain.
 CC It is also expressed in several other tissues. The 15625 receptor
 CC protein may be useful for producing antibodies which can be used to
 CC detect the presence of the receptor protein. The 15625 receptor protein
 CC polynucleotides are useful for generating probes, primers and antisense
 CC constructs. The polynucleotides encoding the 15625 receptor proteins can
 CC also be inserted into vectors to be used in gene therapy. The disorder
 CC that may be treated using the 15625 receptor protein polynucleotides and
 CC polypeptides include anaemia, neutropenia and thrombocytopenia.
 XX SQ Sequence 2286 BP; 786 A; 432 C; 353 G; 715 T; 0 other;
 Query Match 97.5%; Score 1156.8; DB 21; Length 2286;
 Best Local Similarity 99.8%; Pred. No. 1.4e-296;
 Matches 1158; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 388 TTTTGTGTGTCAGTTACCTCCGTCATATTTTATTTTACAAATGTATATCAGTATTTCAAT 447
 DB 542 TTTTGTGTGTCAGTTACCTCCGTCATATTTTATTTTACAAATGTATATCAGTATTTCAAT 601
 QY 448 CCTGGGACTGATTAATATGATCGCTACCAAGACACACCGCCATTAAACATCCAA 507
 DB 602 CCTGGGACTGATTAATATGATCGCTACCAAGACACACCGCCATTAAACATCCAA 661
 QY 508 CCCCCAAATGCTGGGGGGCTAAGATCTCTGTCGATCCTGGGATTCATCTCTT 567
 DB 662 CCCCCAAATGCTGGGGGGCTAAGATCTCTGTCGATCCTGGGATTCATCTCTT 721
 QY 568 ACTCTCTTTGGCTTAACATGATTTGACCAACAGGACCGGACAGCAAGAAATGTGAAGA 627
 DB 722 ACTCTCTTTGGCTTAACATGATTTGACCAACAGGACCGGACAGCAAGAAATGTGAAGA 781
 QY 628 ATGCTCTTTCCCTTAATATGAGATTCGGTCTAGCTGCGCATGAATATGTAATATCATCTG 687
 DB 782 ATGCTCTTTCCCTTAATATGAGATTCGGTCTAGCTGCGCATGAATATGTAATATCATCTG 841
 QY 688 TCAAGTCATTTTCTGGATTAATTTCTTAATGTTATGTTATGTTATACATCATTCACAA 747
 DB 842 TCAAGTCATTTTCTGGATTAATTTCTTAATGTTATGTTATGTTATACATCATTCACAA 901
 QY 748 AGAAGTCACGCGTCATACGTAAAGACGAGGGGTAGTAAAGTCCCGAGAAAAAGT 807
 DB 902 AGAAGTCACGCGTCATACGTAAAGACGAGGGGTAGTAAAGTCCCGAGAAAAAGT 961
 QY 808 GAAAGTCAAAGTTTCATTAATATTCGTCGTATCTTTATTTGTTTGTCTTCATTT 867
 DB 962 GAAAGTCAAAGTTTCATTAATATTCGTCGTATCTTTATTTGTTTGTCTTCATTT 1021
 QY 868 TGCCCGAATTCCTTACACCTGAGCCAAACCGGGATGCTTGTGCTGACGTGGAAGA 927
 DB 1022 TGCCCGAATTCCTTACACCTGAGCCAAACCGGGATGCTTGTGCTGACGTGGAAGA 1081
 QY 928 TACTGTCTTAATGTGAAGAAGACGACCTGTGTTAACTTCCTTAATATGATGCTGGA 987
 DB 1082 TACTGTCTTAATGTGAAGAAGACGACCTGTGTTAACTTCCTTAATATGATGCTGGA 1141
 QY 988 TCCGTCATCTATTTTCTTCTTGTGAAGTCCCTGCAAAATTCCTTGAATATGATGCTGGA 1047
 DB 1142 TCCGTCATCTATTTTCTTCTTGTGAAGTCCCTGCAAAATTCCTTGAATATGATGCTGGA 1201
 QY 1048 GTGCCCAATTCCTGCAACATCTCTGTCAGGACATAGGAAGAAAGACAGATGTGG 1107
 DB 1202 GTGCCCAATTCCTGCAACATCTCTGTCAGGACATAGGAAGAAAGACAGATGTGG 1261
 QY 1108 TGACCCAAATGAGAGACTCCATGTAAACAAATTAAGTAAGAAATATTTCAATCTCTT 1167
 DB 1262 TGACCCAAATGAGAGACTCCATGTAAACAAATTAAGTAAGAAATATTTCAATCTCTT 1321
 QY 1168 TGTGTTCAAGAACTCGTTAAA 1187
 DB 1322 TGTGTTCAAGAACTCGTTAAA 1341

RESULT 5
 AAS98079 standard; DNA; 1332 BP.
 AAS98079;
 12-MAR-2002 (first entry)

Human DNA for potential G protein-coupled receptor #36.
 Human; G protein-coupled receptor; GPCR; ds; GAL4; galatin receptor;
 Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
 atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
 chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
 depression; epilepsy; macular degeneration; lymphoma; melanoma;

KM	multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KM	psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;
KM	tuberculosis; cognition disorder; memory disorder; anorexia;
KM	hormonal release disorder; cardiovascular activity disorder;
KM	pain perception disorder; obesity; diabetes; obesity;
KM	diabetes; hyperlipidemia; stroke; gene therapy.
XX	
OS	Homo sapiens.
PN	WO200185791-A1.
PD	15-NOV-2001.
XX	
PF	11-MAY-2001; 2001WO-US15332.
XX	
PR	11-MAY-2000; 2000US-203217P.
PR	18-MAY-2000; 2000US-205945P.
XX	
PA	(LIFE-) LIFESPAN BIOSCIENCES INC.
PI	Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;
XX	WPI, 2002-066595/09.
XX	
PT	Novel G protein-coupled receptor polypeptides including galanin
PT	receptor polypeptides useful for identifying modulators that are useful
PT	for treating Alzheimer's disease, psoriasis, melanoma, multiple
PT	sclerosis, stroke
XX	
PS	Disclosure; Page 138-139; 144pp; English.
CC	
CC	The invention relates to an isolated polypeptide encoded by a
CC	nucleic acid molecule that is at least 80% identical to the G
CC	protein-coupled (GPCR) polynucleotides included in the specification.
CC	Also included are probes based on the GPCR sequences (including
CC	antisense probes), a host cell comprising an expression vector comprising
CC	the GPCR sequence, antibodies raised against the polypeptides,
CC	and methods of identifying modulators of the polypeptides. The
CC	polypeptides are useful for identifying modulator compounds which
CC	function as modulators, activators, repressors, agonists or antagonists
CC	of the novel GPCR polypeptides including the GAL4 polypeptide. The
CC	antibodies and nucleic acid probes as described above can be used to
CC	detect the presence of the polypeptides and nucleic acids and are used to
CC	diagnose a variety of diseases or disorders in which GPCRs are involved
CC	e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,
CC	atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,
CC	chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,
CC	depression, epilepsy, macular degeneration, lymphoma, melanoma,
CC	multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,
CC	psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis, the
CC	tuberculosis and many other diseases listed in the specification. The
CC	probes and antibodies are also useful for diagnosing cognition and memory
CC	disorders, anorexia, hormonal release disorders, cardiovascular activity
CC	disorders, pain perception disorders, obesity, diabetes, Alzheimer's
CC	disease. Preferably, compounds that decrease or increase
CC	the expression of galanin receptor (GAL4) can be used to treat obesity,
CC	diabetes, hyperlipidemia and stroke. The GPCR nucleic acid is
CC	useful for treating the above mentioned disorders by gene therapy
CC	techniques. The present sequence is a novel GPCR polynucleotide of the
CC	invention.
XX	
SO	Sequence 1332 BP; 416 A; 289 C; 226 G; 401 T; 0 other;
QY	
Query Match	97.2%; Score 1153.4; DB 24; Length 1332;
Best Local Similarity	99.9%; Pred. No. 8-8e-296;
Matches 1154; Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	33 GCGTCGCAATAGTACTACTGATGATCATTCAACCCCTCCAGATCAACAGTTATCAG 92
DB	1 GCGTCGCAATAGTACTACTGATGATCATTCAACCCCTCCAGATCAACAGTTATCAG 60
QY	93 GTACCAACACAAAGTAATGCAGCCGTCGACATCTCACCCTCTGCGCTGGGAAACCCAGTC 152

D	b	61	GTACCAACAAGAAATGCAAGACCGGTGCACAACCTCACTTCCTGCGCCTGGGAAACACCAAGTC	120
Q	y	153	TGTGCACACGAGACTATCAAAATATCCAGAGGCTCTCTTCCACAGTCTCACTGTCTCT	212
D	b	121	TGTGCACACGAGACTATCAAAATATCCAGAGGCTCTCTTCCACAGTCTCACTGTCTCT	180
Q	y	213	TTTTTGTGGACTTATCACAATAGGCTGGCGATGAGATTTTCTTTCAAAATCCGAGAT	272
D	b	181	TTTTTGTGGACTTATCACAATAGGCTGGCGATGAGATTTTCTTTCAAAATCCGAGAT	240
Q	y	273	AATCAACTTATATTTTCTTAAAGAACACTCATTTTCTGATCTTCTCATGATTTCTGA	332
D	b	241	AATCAACTTATATTTTCTTAAAGAACACTCATTTTCTGATCTTCTCATGATTTCTGA	300
Q	y	333	CTTTTCCATTCAAAATTTCTTAGTAGATCCCAACTGGGAACAGACACATGAGACCTTTTG	392
D	b	301	CTTTTCCATTCAAAATTTCTTAGTAGATCCCAACTGGGAACAGACACATGAGACCTTTTG	360
Q	y	393	TGTGTCAAGTAACTCCGTCATATTTTATTTTCAATGTATATCAATATTTCAATTCCTGG	452
D	b	361	TGTGTCAAGTAACTCCGTCATATTTTATTTTCAATGTATATCAATATTTCAATTCCTGG	420
Q	y	453	GACTGATATCAATGATGATCGCTACCAGAAACACACAGGCCATTTAAACATCCAACCCCA	512
D	b	421	GACTGATATCAATGATGATCGCTACCAGAAACACACAGGCCATTTAAACATCCAACCCCA	480
Q	y	513	AAATATCTTTGGGGGCTTAAGATTTCTCTCTTGTCTACATCGGGCATTCATGTTCTTACTCT	572
D	b	481	AAATATCTTTGGGGGCTTAAGATTTCTCTCTTGTCTACATCGGGCATTCATGTTCTTACTCT	540
Q	y	573	CTTTGCTCAATCATGATTTCTGACCACAGCGACCGAGACACAAGATGTGAAGAAATGCT	632
D	b	541	CTTTGCTCAATCATGATTTCTGACCACAGCGACCGAGACACAAGATGTGAAGAAATGCT	600
Q	y	633	CTTTTCCCTTAATTCGAGATTTGGGTCTAGTCTGGCATTAATTAATTAATTCATCTGTCAAG	692
D	b	601	CTTTTCCCTTAATTCGAGATTTGGGTCTAGTCTGGCATTAATTAATTAATTCATCTGTCAAG	660
Q	y	693	TCATTTTCTGGATTAATTTCTTAATTAATGTTATTTGATGTTATTCACATTTACAAAAGAC	752
D	b	661	TCATTTTCTGGATTAATTTCTTAATTAATGTTATTTGATGTTATTCACATTTACAAAAGAC	720
Q	y	753	TGTACCGGTCAATACGATGAAGACGAGGGGTGATGAATGCCCCAGAAAAAGGTGAACG	812
D	b	721	TGTACCGGTCAATACGATGAAGACGAGGGGTGATGAATGCCCCAGAAAAAGGTGAACG	780
Q	y	813	TCAAGTTTTCATTAATCATATGTCGTATTCCTTATTTGTTTGTCTTCCATTTTGGCC	872
D	b	781	TCAAGTTTTCATTAATCATATGTCGTATTCCTTATTTGTTTGTCTTCCATTTTGGCC	840
Q	y	873	GAAATCCCTTAACACCTGAGCCCAAAACCCGGAGTCTTTTACATGCACTGCTGAAATATCNC	932
D	b	841	GAAATCCCTTAACACCTGAGCCCAAAACCCGGAGTCTTTTACATGCACTGCTGAAATATCNC	900
Q	y	933	TGTTCTATGTGAAGAGAGACACTGTGTGTTAACTTCTTAAATGCATGCCGTGATCCGT	992
D	b	901	TGTTCTATGTGAAGAGAGACACTGTGTGTTAACTTCTTAAATGCATGCCGTGATCCGT	960
Q	y	993	TGCATCTATTTTTTCTTGAAGTCCCTTGAGAAATTCCTGTAATGTATGTGTAAGTGTGC	1052
D	b	961	TGCATCTATTTTTTCTTGAAGTCCCTTGAGAAATTCCTGTAATGTATGTGTAAGTGTGC	1020
Q	y	1053	CCAATTTCTGCAACATCTCTGCCAGAGACAATGTGAAGAAAAAGACAGATGTGTGTGAC	1112
D	b	1021	CCAATTTCTGCAACATCTCTGCCAGAGACAATGTGAAGAAAAAGACAGATGTGTGTGAC	1080
Q	y	1113	CAAAATGAAGAGACTCAATGTAAACAAATTAATTAAGAAATATTTCAATCTCTTGTGT	1172
D	b	1081	CAAAATGAAGAGACTCAATGTAAACAAATTAATTAAGAAATATTTCAATCTCTTGTGT	1140
Q	y	1173	TCGAGACTCGTTAAA 1187	
D	b	1141	TCAGAACTCGTTAAA 1155	

RESULT 6
 AAD08695
 ID AAD08695 standard; cDNA; 1330 BP.
 XX
 AC AAD08695;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human P2-purine receptor subtype, P2Y12 cDNA.
 XX
 KW Human; P2-purine receptor; P2Y12; cardiant; vasotropic; thrombolytic;
 KW cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina
 KW myocardial infarction; ischemic attack; preclampsia; bleeding disorder;
 KW carotid endarterectomy; vascular graft surgery; brain disorder; migraine;
 KW vascular injury; schizophrania; eating disorder; depression; angioplasty;
 KW peripheral vascular disease; platelet aggregation; restenotic; embolism;
 KW thrombocytopaenic purpura; stroke; pertussis toxin sensitive G protein;
 KW G1; disseminated intravascular coagulation; thrombosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 73..1101
 FT /tag= a
 FT /product= "Human P2-purine receptor, P2Y12"
 XX
 PN WO200146454-A1.
 XX
 PD 28-JUN-2001.
 XX
 PE 26-DEC-2000; 2000MO-US34998.
 XX
 PR 23-DEC-1999; 9905-0171622.
 XX
 PA (COR-) COR THERAPEUTICS INC.
 XX
 PI Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;
 PI Hollpeter G;
 XX
 DR WPI: 2001-418082/44.
 XX
 PS P-PSDB: AAE04386.
 XX
 PT Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful
 PT for identifying binding partners and for diagnostic applications -
 XX
 PS Example 1; Page 85-87; 108pp: English.
 XX
 CC The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed
 CC as P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is
 CC the subtype of P2-purine receptor. The P2Y12 receptor is expressed
 CC selectively in the platelets and brain, and couples to a pertussis toxin
 CC sensitive G protein (G1). P2Y12 receptor is a G protein-coupled receptor
 CC that responds to ADP. The invention also relates to a method for
 CC identifying an agent which is useful for modulating acute myocardial
 CC infarction, unstable angina, chronic stable angina, transient ischaemic
 CC attacks, strokes, peripheral vascular disease, preclampsia, deep venous
 CC thrombosis, embolism, disseminated intravascular coagulation, thrombotic
 CC thrombocytopaenic purpura or a bleeding disorder; thrombotic and
 CC restenotic complications following angioplasty, carotid endarterectomy,
 CC post CABG (coronary artery bypass graft) surgery, vascular graft surgery,
 CC stent placements or insertion of endovascular devices and prostheses.
 CC P2Y12 receptor is useful for identifying binding partners and for
 CC diagnostic applications. P2Y12 receptor provides targets for screening
 CC synthetic small molecules and combinatorial or naturally occurring
 CC compound libraries to regulate platelet aggregation, vascular injury, or
 CC disease as well as schizophrania, eating disorders, depression, migraine
 CC and other brain disorders.
 CC
 CC The present cDNA sequence encodes human P2-
 CC purinergic receptor subtype, referred as P2Y12 receptor.
 XX

Query Match	97.0%;	Score 151.4;	DB 22;	Length 1330;	
Best Local Similarity	99.9%;	Pred. No. 3e-295;			
Matches 1152;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	35	CTGCATTAACACTCTACTACTAGTATGATACATTCCTCAACCCCTCCAGAAATCAACAGTTATCAGGT	94		
Db	1	CTGGCAATAACACTCTACTACTAGTATGATACATTCAAACCCCTCCAGAAATCAACAGTTATCAGGT	60		
QY	95	AACCAACAAGAAATGCAGACCGCTGCACAAATCTCACCTCTGGCCCTGGGAACCAACAGCTGTG	154		
Db	61	AACCAACAAGAAATGCAGACCGCTGCACAAACCTCACCTCTGGCCCTGGGAACCAACAGCTGTG	120		
QY	155	TGCACACAGAGACTACAAAAATCACCCAGAGGTCCTTTCCCGCATGCTGTACACCTGCTGGTT	214		
Db	121	TGCACACAGAGACTACAAAAATCACCCAGAGGTCCTTTCCCGCATGCTGTACACCTGCTGGTT	180		
QY	215	TTTGTGGACCTTATCACAANAATGGCTGGCGATGAGAGATTTCTTCAAAATCCGAGTAA	274		
Db	181	TTTGTGGACCTTATCACAANAATGGCTGGCGATGAGAGATTTCTTCAAAATCCGAGTAA	240		
QY	275	TCAAACTTATATTTTCTTAAAGAACAGTCAATTTCTGATCTTCTCATGATTTCTGACT	334		
Db	241	TCAAACTTATATTTTCTTAAAGAACAGTCAATTTCTGATCTTCTCATGATTTCTGACT	300		
QY	335	TTTTCAATTCAAAATTTCTTAGTATGACCAAACTGGGAAACAGAGCCACTGAGAACTTTTGTG	394		
Db	301	TTTTCAATTCAAAATTTCTTAGTATGACCAAACTGGGAAACAGAGCCACTGAGAACTTTTGTG	360		
QY	395	TGTCAAGTTACTCCGCTCATATTTTATTTACAAATGTATATCATATTTTCTTCTGGGA	454		
Db	361	TGTCAAGTTACTCCGCTCATATTTTATTTACAAATGTATATCATATTTTCTTCTGGGA	420		
QY	455	CTGATAACATCATGATCGCAACAGCAAGACACACAGCCACTTTTAAAAATCTCAACCCCA	514		
Db	421	CTGATAACATCATGATCGCTACCAAGAAAGACACACAGCCATTTTAAAAATCTCAACCCCA	480		
QY	515	AATCTCTTGGGGGGCTAAGATTTCTCTCTGTTCATCTGGGCAATTCATGTTCCTACTCTCT	574		
Db	481	AATCTCTTGGGGGGCTAAGATTTCTCTCTGTTCATCTGGGCAATTCATGTTCCTACTCTCT	540		
QY	575	TTGGCTAACATATTTCTGCAACAAAGGGCAGCGAGAGACAAAGATGTGAAGAAATGCTCT	634		
Db	541	TTGGCTAACATATTTCTGCAACAAAGGGCAGCGAGAGACAAAGATGTGAAGAAATGCTCT	600		
QY	635	TTTCTTAAATCAGAGTTTCGCTAGTCTGCGCATGAAATAGTAAATTTACATCTGCTCAAGTC	694		
Db	601	TTTCTTAAATCAGAGTTTCGCTAGTCTGCGCATGAAATAGTAAATTTACATCTGCTCAAGTC	660		
QY	695	ATTTTCTGGATTAATTTCTTAATTTGTTATTTGATATGTTATACACTCATTTCAAAAAGACATG	754		
Db	661	ATTTTCTGGATTAATTTCTTAATTTGTTATTTGATATGTTATACACTCATTTCAAAAAGACATG	720		
QY	755	TACCGGTCTACGTAAAGACGAGGGGTGTAAGTAAAGTCCCGCAGAAAAAGGTGAACGTC	814		
Db	721	TACCGGTCTACGTAAAGACGAGGGGTGTAAGTAAAGTCCCGCAGAAAAAGGTGAACGTC	780		
QY	815	AAAGTTTTCATTATCATTTGCTGTATTTCTTTATTTGTTTGTTCCTTCCATTTTGCCCGA	874		
Db	781	AAAGTTTTCATTATCATTTGCTGTATTTCTTTATTTGTTTGTTCCTTCCATTTTGCCCGA	840		
QY	875	ATTTCCTTACCGCTGAGCCAAACCGGGATGTCTTGATGTGACATGCTGATGCTGAAATATCTGTG	934		
Db	841	ATTTCCTTACCGCTGAGCCAAACCGGGATGTCTTGATGTGACATGCTGATGCTGAAATATCTGTG	900		
QY	935	TTTCTATGTGAAGAGAGACACTGTGTGTTAACCTTCTTTAAATGCAATGCTGTGATCCGTC	994		
Db	901	TTTCTATGTGAAGAGAGACACTGTGTGTTAACCTTCTTTAAATGCAATGCTGTGATCCGTC	960		
QY	995	ATCTATATTTTCTCTTTTGCAAGTCCTTCAGAAATTCCTTATATAGTATGCTGGAAGTCCCC	1054		
Db	961	ATCTATATTTTCTCTTTTGCAAGTCCTTCAGAAATTCCTTATATAGTATGCTGGAAGTCCCC	1020		
QY	1055	AATTTGTGCACATCTCTGTGCCAGGACATAGAAAAAAGAACAGAGATGCTGTGACCA	1114		

```

|||||
Db 1021 AATCTGCAACATCTCTGTCTCCAGACATATGAAAAAGACGATGGTGACCA 1080
QY 1115 AATGAGAGACTCCAAATGTAAACAATTAACTAAGAGAAATTTCAATCTCTTGTGTTTC
Db 1081 AATGAGAGACTCCAAATGTAAACAATTAACTAAGAGAAATTTCAATCTCTTGTGTTTC 1140
QY 1175 AGAAGCTGCTTAAA 1187
Db 1141 AGAAGCTGCTTAAA 1153

RESULT 7
AAD08700
ID AAD08700 standard; cDNA; 1328 BP.
AC AAD08700;
XX
XX
XX 04-SEP-2001 (first entry)
XX
XX Human P2-purineric receptor, P2Y12 truncated allelic variant cDNA.
XX
XX Human; P2-purineric receptor; P2Y12; cardiant; vasotrophic; thrombolytic;
XX cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina;
XX myocardial infarction; ischaemic attack; preeclampsia; bleeding disorder;
XX carotid endarterectomy; vascular graft surgery; brain disorder; migraine;
XX vascular injury; schizophrenia; eating disorder; depression; angioplasty;
XX peripheral vascular disease; platelet aggregation; restenosis; embolism;
XX thrombocytopenic purpura; stroke; pertussis toxin-sensitive G protein;
XX G1; disseminated intravascular coagulation; thrombosis; mutant; mutein;
XX variant; ss.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 73..876
XX FT /*tag= a
XX FT /product= "Human P2-purineric receptor P2Y12 mutant"
XX FT
XX PN WC200146454-A1.
XX
XX PD 28-JUN-2001.
XX
XX PF 26-DEC-2000; 2000MO-US34998.
XX
XX PR 23-DEC-1999; 9905-0171622.
XX
XX PA (CORI-) COR THERAPEUTICS INC.
XX
XX PI Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;
XX PI Holloper G;
XX PI WPI: 2001-418082/44.
XX DR P-PSDB; AAE04387.
XX
XX PT Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful
XX PT for identifying binding partners and for diagnostic applications
XX PT
XX PS Claim 19; Page 90-91; 108bp; English.
XX
XX The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed
XX as P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is
XX the subtype of P2-purineric receptor. The P2Y12 receptor is expressed
XX selectively in the platelets and brain, and couples to a pertussis toxin-
XX sensitive G protein (G1). P2Y12 receptor is a G protein-coupled receptor
XX that responds to ADP. The invention also relates to a method for
XX identifying an agent which is useful for modulating acute myocardial
XX infarction, unstable angina, chronic stable angina, transient ischaemic
XX attacks, strokes, peripheral vascular disease, preeclampsia, deep venous
XX thrombosis, embolism, disseminated intravascular coagulation, thrombotic
XX thrombocytopenic purpura or a bleeding disorder; thrombotic and
XX restenotic complications following angioplasty, carotid endarterectomy,

```

```

CC post CABG (coronary artery bypass graft) surgery, vascular graft surgery,
CC stent placements or insertion of endovascular devices and prostheses.
CC P2Y12 receptor is useful for identifying binding partners and for
CC diagnostic applications. P2Y12 receptor provides targets for screening
CC synthetic small molecules and combinatorial or naturally occurring
CC compound libraries to regulate platelet aggregation, vascular injury, or
CC disease as well as schizophrenia, eating disorders, depression, migraine
CC and other brain disorders. The present cDNA sequence encodes human P2-
CC purineric receptor subtype, P2Y12 truncated allelic variant. This
CC variant cDNA is obtained by deleting two bases 'CA' from position 769-790
CC of the wildtype P2-purineric receptor, P2Y12 cDNA.
XX
SQ Sequence 1328 BP; 415 A; 288 C; 224 G; 401 T; 0 other;
Query Match 95.8%; Score 1137.4; DB 22; Length 1328;
Best Local Similarity 99.7%; Pred. No. 1,5e-291;
Matches 1150; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 35 CTGCATTAATCTACTTACTTGTGATGATTCATTAACCTCCAGAAATCAAGTTATTCAGGT 94
Db 1 CTGCATTAATCTACTTACTTGTGATGATTCATTAACCTCCAGAAATCAAGTTATTCAGGT 60
QY 95 AACCAACAGAAATGCAAGCGGTGACAAATCTCACCTGCGGCTGGGACACCAAGTCG 154
Db 61 AACCAACAGAAATGCAAGCGGTGACAAATCTCACCTGCGGCTGGGACACCAAGTCG 120
QY 155 TGCACAGAGACTACAAATACCAACAGGCTCTTCCACCTGCTACACTGTCCTGTTT 214
Db 121 TGCACAGAGACTACAAATACCAACAGGCTCTTCCACCTGCTACACTGTCCTGTTT 180
QY 215 TTTGTGTGACTTATCACAAATGGCTGGCATGAGATTTTCTTCAATCCGAGTAAA 274
Db 181 TTTGTGTGACTTATCACAAATGGCTGGCATGAGATTTTCTTCAATCCGAGTAAA 240
QY 275 TCAACCTTTATTTATTTTCTTAAAGAACAGTCAATTCGATCTTCATGATTTGACT 334
Db 241 TCAACCTTTATTTATTTTCTTAAAGAACAGTCAATTCGATCTTCATGATTTGACT 300
QY 335 TTTTCATTTCAAAATTTCTTACGTATGCCAAATCGGGAACAGACCACTGAGACCTTTTGTG 394
Db 301 TTTTCATTTCAAAATTTCTTACGTATGCCAAATCGGGAACAGACCACTGAGACCTTTTGTG 360
QY 395 TGTCAATTTACCTCCGATATTTTATTTACAAATGATATTCAGTATTTCTTGGGA 434
Db 361 TGTCAATTTACCTCCGATATTTTATTTACAAATGATATTCAGTATTTCTTGGGA 420
QY 455 CTGATACTATGATGCTACGACCAAGAGACCAAGGCAATTTAAACATCAACCCAAA 514
Db 421 CTGATACTATGATGCTACGACCAAGAGACCAAGGCAATTTAAACATCAACCCAAA 480
QY 515 AATCTCTTGGGGGCTAAGATTTCTCTGTTGTGATCTGGGCATTCATGTTCTTACTCTCT 574
Db 481 AATCTCTTGGGGGCTAAGATTTCTCTGTTGTGATCTGGGCATTCATGTTCTTACTCTCT 540
QY 575 TTGGCTAACATGATTTCTGACCAACAGGACCCAGAGACAAAGATGGAAGAAATGCTCT 634
Db 541 TTGGCTAACATGATTTCTGACCAACAGGACCCAGAGACAAAGATGGAAGAAATGCTCT 600
QY 635 TTCCCTTAATCAAGATTCGCTAGTCTGATGCAATTAATTAATTAATCAATCTGCAAGTC 694
Db 601 TTCCCTTAATCAAGATTCGCTAGTCTGATGCAATTAATTAATTAATCAATCTGCAAGTC 660
QY 695 ATTTTCTGATTAATTTCTTAATTTGTTATGTTATACACTCAATTCACAAAGAACTG 754
Db 661 ATTTTCTGATTAATTTCTTAATTTGTTATGTTATACACTCAATTCACAAAGAACTG 720
QY 755 TACCGGTCATACCTAAGACAGAGGGGTGATGTAAGTCCCAAGAAAAGTGAACGTC 814
Db 721 TACCGGTCATACCTAAGACAGAGGGGTGATGTAAGTCCCAAGAAAAGTGAACGTC 780
QY 815 AAGGTTTTCATTTCAATGCTGATTTCTTATTTGTTTGTCTTCCATTTTGGCCGA 874
Db 781 AAGGTTT--TTATCATGTGCTGATTTCTTATTTGTTTGTCTTCCATTTTGGCCGA 838

```

QY	875	ATTCCTTACACCTCGAACCCAAACCCGGGATGTCCTTGGACTGCAGCTGTGAAAATACCTGTG	934
Db	839	ATTCTCTTACACCCCGAGACCAAAACCCGGAGATGTCCTTGGACTGCAGCTGTGAAAATACCTGTG	898
QY	935	TTCTATGTGSAAGAGACACTGTGTGGTTAACTTCTTAAATGCATGCTGGATTCGGTTC	994
Db	899	TTCTATGTGSAAGAGACACTGTGTGGTTAACTTCTTAAATGCATGCTGGATTCGGTTC	958
QY	995	ATCTATTTTTTCCCTTTCGACAGTCTCTTCAGAAATTCCTGATTAAGATGCTGAAGTGCCCC	1054
Db	959	ATCTATTTTTTCCCTTTCGACAGTCTCTTCAGAAATTCCTGATTAAGATGCTGAAGTGCCCC	1018
QY	1055	AATTCCTGCACACATCTCTGTCCACAGCAATAGGAAAAAAGAACAGAGATGTTGTGACCCA	1114
Db	1019	AATTCCTGCACACATCTCTGTCCACAGCAATAGGAAAAAAGAACAGAGATGTTGTGACCCA	1078
QY	1115	AATGAGAGACTCCATGTAACAAATTAATCTAAGGAAATATTTCAATCTCTTTGTGTTG	1174
Db	1079	AATGAGAGACTCCAAATGTAACAAATTAATCTAAGGAAATATTTCAATCTCTTTGTGTTG	1138
QY	1175	AGAACTCGTTAAA	1187
Db	1139	AGAACTCGTTAAA	1151

RESULT 8
AAA27127
ID AAA27127 standard; DNA; 2272 BP.

AA
AC AAA27127;

DT 21-AUG-2000 (first entry)

Macaque ortholog of human 15625 receptor protein cDNA.

KW Human; G-protein-coupled receptor; GPCR; 15625 receptor protein;

KW red cell; thymus; B-cell; pancreas; disorder; chromosome 3;

XX

XX

FT CDS

/product=	"15625 receptor protein"
FT	

PN WO200028028-A1

PD 18-MAY-2000.

PF 05-NOV-1999; 99WO-US25956.

PR 06-NOV-1998; 98US-0187134.

XX

XX

XX

DR P-PSDB; AAY94445.

PT Identifying an agent modulating the level or activity of G-protein

PT having disorders such as anaemia, neutropenia and thrombocytopenia

PS Disclosure; Page 92-93; 97pp; English.

CC The 15625 receptor protein is a novel G-coupled protein receptor (GPCR).

CC with sequences homologous to other GPCRs. The 15625 receptor protein is

expressed in the glial cells of the brain. It is also expressed in several other tissues. The 15625 receptor protein may be useful for producing antibodies which can be used to detect the presence of the receptor protein. The 15625 receptor protein polynucleotides are useful for generating probes, primers and antisense constructs. The polynucleotides encoding the 15625 receptor proteins can also be inserted into vectors to be used in gene therapy. The disorder that may be treated using the 15625 receptor protein polynucleotides and polypeptides include anaemia, neutropenia and thrombocytopenia. The present sequence is the macaque ortholog of the human 15625 receptor protein cDNA.

Query Match	93.48;	Score 1108.8;	DB 21;	Length 2272;
Best Local Similarity	97.28;	Pred. No. 7.4e-284;		
Matches 1128; Conservative	0;	Mismatches 32;	Indels 0;	Gaps 0

OY	28	CTGACGGCTGCATATACCTACTCTACTTCTGCGATTCATTCAAACCCTCCAGAAATCAACAGTT	87
Db	160	CTGCACGTTGCATATACCTACTATTTTACTGGATTCATTTCAATCTCTCAGAAATCAACGGTT	219
OY	88	ATCAGGTAACCAACAAGAAATGCACGCCCTGCACAACTTCACCTCTCCGCCCTGGGAACAC	147
Db	220	ATCAGGTAACCAACAAGAAATGCACGCCATTCGACAACTCCAGCTCGCGCTGGGAACAC	279
OY	148	CAGCTCTGCACCCAGAGACTACAAAATCACACCAGTCTCTTCCCACTGCTTACACTGT	207
Db	280	CAGTCTGTGCACACAGAGACTACAAAATCACACCAGTCTCTTCCCACTGCTTACACTGT	339
OY	208	CCGTGTTTTTGGTTGGACTATTCACAAATGAGCCCGGGGATGAGGATTTTCTTTCAAAATCCG	267
Db	340	CCTGTTTTTGTGTTGGACTATTCACAAATGAGCTGGGATGAGGATTTTCTTTCAAAATCCG	399
OY	268	GAGTAATCAACCTTTATTTATTTTCTTAAGAACACAGTCAATTTTCAGATCTTCTCATGAT	327
Db	400	GAGTAATCAACCTTTATTTATTTTCTTTAAGAACACAGTCAATTTTCGATCTTCTCATGAT	459
OY	328	TCTGACTTTTCCATTCAAAATTTCTTAGTGTATGCCAAACTGGGAACAGACACTGAGAAC	387
Db	460	TCTGACTTTTCCATTCAAAATTTCTTAGTGTATGGCAAACTGGGAACAGACACTGAGAAC	519
OY	388	TTTTTGTTGTCAAGTTACCTCCGCTCATATTTTATTTTTCACAAATGTATATGATTTTTCATT	447
Db	520	TTTTTGTTGTCAAGTTACTCCGCTCATATTTTATTTTTCACAAATGTATATGATTTTTCATT	579
OY	448	CTGTGGACTGATTAACATAGATCGATCGTACAGAAAGACCACAGGCCATTTAAAAATCCAA	507
Db	580	CTGTGGACTGATTAACATAGATCGATCGTACAGAAAGACCACAGGCCATTTAAAAATCCAA	639
OY	508	CCCCAAATATCTTTGGGGGCTTAAGATTTCTCTCTGTGTTCATCTGGGCATTCATGTTCTT	567
Db	640	CCCCAAATATCTTTGGGGGCTTAAGATTTCTCTCTGTGTTCATCTGGGCATTCATGTTCTT	699
OY	568	ACTCTCTTGGCTTAACATATATCTTGACCAACAGCAGCAGCAGAGAGAAGAATGTGAAGAA	627
Db	700	ACTCTCTTGGCTTAACATATATCTTGACTTAACAGCGCCGCCAAGAGACAAGAATGTGAAGAA	759
OY	628	ATGCTCTTCCCTTAATTCAGATGGCGTCTAGTCTGGCATGAATAATGTAATATCATCTG	667
Db	760	ATGCTCTTCCCTTAATTCAGATGGCGCTAGTCTGGCATGAATAATGTAATATCATCTG	819
OY	688	TCAAGCATTTTCTGGAATTAATTTCTTAATGTATTTGATATGTTATTAACCTCAATTCAAA	747
Db	820	TCAAGCATTTTCTGGAATTAATTTCTTAATGTATTTGATATGTTATTAACCTCAATTCAAA	879
OY	748	AGNACTGTACCGGTCTATAGTAAAGAACGAGGGGTGTAGTAAAGTCCCCAGAAAAAGGT	807
Db	880	AGNACTGTACCGGTCTATATGTAAAGAACAGGGGTGTAGTAAAGTCCCCAGAAAAAGGT	939
OY	808	GAACTGCAAAAGTTTATATATCATTCGCTGATTTCTTTATTTGTTTGTTCCTTTCATTT	867
Db	940	GAACTGCAAAAGTTTATATATCATTCGCTGATTTCTTTATTTGTTTGTTCCTTTCATTT	999

QY 868 TGGCCCAATTCCTTACACCTGAGCCAAACCGGGATGCTTTGACTGCACCTCTGAAAA 927
 |||||||
 Db 1000 TGGCCCAATTCCTTATACCTGAGCCAAACCGGGATGCTTTGACTGCACCTCTGAAAA 1059
 QY 928 TACTGTTCTATGTGAAGAGAGACACTGTGTGTTAACTTCCCTTAATGATGATGCTGGA 987
 |||||||
 Db 1060 TACTGTTCTATGTGAAGAGAGACTGTGTGTTAACTTCCCTTAATGATGATGCTGGA 1119
 QY 988 TCCGTTACATCTATTTTCTTTGCAAGTCTTCAGAAATCTCTGATTAATGATGCTGAA 1047
 |||||||
 Db 1120 TCCGTTACATCTATTTTCTTTGCAAGTCTTCAGAAATCTCTGATTAATGATGCTGAA 1179
 QY 1048 GTGCCCCAATTCGCAACATCTCTGTCGCCAGACATAGAGAAAAAGACAGATGCTGG 1107
 |||||||
 Db 1180 GTGCCCCAATTCGCAACATCTCTGTCGCCAGACATAGAGAAAAAGACAGATGCTGG 1239
 QY 1108 TGACCCCAATGAGAGAGACCTCCAAATGTAACAAATTAATTAATGATGCTGCTT 1167
 |||||||
 Db 1240 TGACCCCAATGAGAGAGACCTCCAAATGTAACAAATTAATTAATGATGCTGCTT 1299
 QY 1168 TGTGTTCAAGAACTGTTAAA 1187
 |||||||
 Db 1300 TGGTTTCAGAACTCATTTAA 1319

RESULT 9
 AA168803
 ID AA168803 standard; DNA; 1101 BP.
 XX

AC AA168803;

DT 24-JAN-2002 (first entry)

XX Human PAFR-3 DNA fragment.

DE PAFR-3; platelet activating factor receptor; human; chromosome 3;

KW thrombocyte activation; hypotension; vascular permeability;

KW bronchoconstriction; transgenic animal; knockout animal; ds.

XX Homo sapiens.

OS Homo sapiens.

XX Key CDS Location/Qualifiers

FT CDS 19..1047

FT /tag= a

FT /product= "PAFR-3"

PN DE10020073-A1.

PD 25-OCT-2001.

XX 22-APR-2000; 2000DE-1020073.

XX 22-APR-2000; 2000DE-1020073.

XX 22-APR-2000; 2000DE-1020073.

XX (BRUE/) BRUESS M.

XX (BOEN/) BOENISCH H.

XX Brues M, Boenisch H;

XX WPI: 2002-027296/04.

XX P-PSDB; AAG80236.

XX New human platelet-activating factor (PAF) receptor-3 gene, useful for

XX diagnosis and treatment of PAF-related diseases

XX Disclosure; Page 4; 6pp; German.

CC (I), and derived (anti)sense oligonucleotides, are useful in the
 CC treatment and diagnosis of (I)-related diseases; for producing
 CC transgenic/knockout animals, and for recombinant expression of the
 CC protein (II) that it encodes. (II) is useful in ligand-binding studies
 CC and screening assays, also for treatment and diagnosis of (II)-related
 CC diseases. This sequence encodes the human leucocyte platelet-activating
 CC factor receptor (PAFR-3) protein described in the invention.
 XX

SQ Sequence 1101 BP; 314 A; 247 C; 195 G; 345 T; 0 other;

Query Match 92.1%; Score 1093.8; DB 24; Length 1101;

Best Local Similarity 99.8%; Pred. No. 5.3e-280;

Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGGTAAACCAACAAAGAAATGACAGCCGCTGACAAATCTACCTGCGCTGGAGAACAG 150
 |||||||
 Db 3 AGGTAAACCAACAAAGAAATGACAGCCGCTGACAAATCTACCTGCGCTGGAGAACAG 62
 QY 151 TCTGTGACACAGAGACTACAAATACACAGTCTCTCCACTGCTACACTGCTCT 210
 |||||||
 Db 63 TCTGTGACACAGAGACTACAAATACACAGTCTCTCCACTGCTACACTGCTCT 122
 QY 211 GTTTTGTGAGACTATACCAAAATGGCTGGGATGAGATTTCTTCAATCCGGAG 270
 |||||||
 Db 123 GTTTTGTGAGACTATACCAAAATGGCTGGGATGAGATTTCTTCAATCCGGAG 182
 QY 271 TAAATCAACTTATTTATTTCTTAAAGACACAGTCAATTCGATCTCATGATCT 330
 |||||||
 Db 183 TAAATCAACTTATTTATTTCTTAAAGACACAGTCAATTCGATCTCATGATCT 242
 QY 331 GACTTTTCCATTAATAATTTTGTGATGCGCAACCTGGAGAACAGACTGAACTTT 390
 |||||||
 Db 243 GACTTTTCCATTAATAATTTTGTGATGCGCAACCTGGAGAACAGACTGAACTTT 302
 QY 391 TGTGTGCAAGTACACCTGCTGATTTTATTTCAAAATGATATGATTTTCAATCT 450
 |||||||
 Db 303 TGTGTGCAAGTACACCTGCTGATTTTATTTCAAAATGATATGATTTTCAATCT 362
 QY 451 GGGAGTGAATACATGATGCTGACAGAGACAGGAGGAGAGAGAGAGAGAGAGAG 510
 |||||||
 Db 363 GGGAGTGAATACATGATGCTGACAGAGAGACAGGAGGAGAGAGAGAGAGAGAG 422
 QY 511 CAAAATCTCTTGGGGCTTAAGATTCTCTGTGTGATCTGGGCTGATCTTCTACT 570
 |||||||
 Db 423 CAAAATCTCTTGGGGCTTAAGATTCTCTGTGTGATCTGGGCTGATCTTCTACT 482
 QY 571 CTCTTGGCTTAATGATGCTGACAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 630
 |||||||
 Db 483 CTCTTGGCTTAATGATGCTGACAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 542
 QY 631 CTCTTGGCTTAATGATGCTGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
 |||||||
 Db 543 CTCTTGGCTTAATGATGCTGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
 QY 691 AGTCATTTTCTGATTAATTTCTTAATTTGTTATGTTATGATGATGATGATGAT 750
 |||||||
 Db 603 AGTCATTTTCTGATTAATTTCTTAATTTGTTATGTTATGATGATGATGATGAT 662
 QY 751 ACTGTACCGGTCAATGATGAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 810
 |||||||
 Db 663 ACTGTACCGGTCAATGATGAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 722
 QY 811 CGTCAAGTTTCTTATCATCTGCTGATTTCTTAATTTGTTTGTGTTTCTTCCATTTG 870
 |||||||
 Db 723 CGTCAAGTTTCTTATCATCTGCTGATTTCTTAATTTGTTTGTGTTTCTTCCATTTG 782
 QY 871 CCGAATTTCTTACACCTGAGCCAAACCGGAGTGTCTTGTGACTGACATGCTGAATAC 930
 |||||||
 Db 783 CCGAATTTCTTACACCTGAGCCAAACCGGAGTGTCTTGTGACTGACATGCTGAATAC 842
 QY 931 TGTGTTCTATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
 |||||||
 Db 843 TGTGTTCTATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902

```
QY 991 GTTCATCTATTTTCTTTCCTTTGCAAGTCCTTCAGAAATTCCTTGATGAATGCTGAAGTG 1050
    |||||||
Db 903 GTTCATCTATTTTCTTTCCTTTGCAAGTCCTTCAGAAATTCCTTGATGAATGCTGAAGTG 962
QY 1051 CCCCAATTCGCAACATCTCTGTCTCCGAGACAAATGAGAAAAAACAAGATGCTGTGA 1110
    |||||||
Db 963 CCCCAATTCGCAACATCTCTGTCTCCGAGACAAATGAGAAAAAACAAGATGCTGTGA 1022
QY 1111 CCCAATGAAGAGACTCCAAATGTAAACAATTAAGAAATATTTCAATCTCTTGT 1170
    |||||||
Db 1023 CCCAATGAAGAGACTCCAAATGTAAACAATTAAGAAATATTTCAATCTCTTGT 1082
QY 1171 GTTCAGAGCTCGTTAA 1187
    |||||||
Db 1083 GTTCAGAGCTCGTTAA 1099

RESULT 10
ABK50286
ID ABK50286 standard; DNA: 1143 BP.
XX
AC ABK50286;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human DNA encoding the ADP-glucose receptor.
XX
Human: ds: gene: ADP-glucose receptor; G protein-coupled receptor;
KM agonist; antagonist; cardiovascular function disorder; vasorelaxation;
KM ischemia; angina pectoris; gastrointestinal disorder; diarrhoea;
KM immune disorder; immunodeficiency disorder; autoimmune disorder;
KM rheumatoid arthritis; bacterial infection; viral infection;
KM fungal infection; protozoal infection; respiratory disorder; asthma;
KM kidney disorder; glomerulonephritis; hepatobiliary disorder; cirrhosis;
KM endocrine disorder; adrenal dysfunction; musculoskeletal disorder;
KM osteoporosis; nervous system disorder; Alzheimer's disease;
KM psychotic disorder; depression; cancer; pain; glycogen storage disease;
KM disorder of body weight; AIDS; acquired immunodeficiency syndrome;
KM chromosome 3; Usher's syndrome type 3.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 52..1080
FT /lag- a
FT /product= "ADP-glucose receptor"
XX
PN W0200224942-A2.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001WO-US29523.
XX
PR 20-SEP-2000; 2000US-234025P.
PR 09-FEB-2001; 2001US-0780576.
XX
PA (REGC ) UNIV CALIFORNIA.
PI Clavelli O, Nothacker H, Wang Z, Reinscheld R;
XX
DR WPI: 2002-383195/41.
DR P-PSDB; AAU80164.
XX
PT Identifying an agonist, antagonist or ligand of an ADP-glucose
PT receptor, for treating cardiovascular, gastrointestinal, kidney,
PT endocrine, immune disorders, and bacterial, viral, protozoal or fungal
PT infections
XX
PS Example 1: Fig 1; 86pp; English.
XX
CC The invention relates to identifying an ADP-glucose (ADP-G) receptor (a
CC G protein-coupled receptor) agonist, antagonist or ligand, comprises
```

```
CC contacting the receptor with one or more candidate compounds so that the
CC receptor produces a G protein-coupled signal in response to ADP-G or
CC selectively binds ADP-G; and identifying the candidate molecule that
CC alters signal production as an agonist, antagonist or binds as a ligand.
CC Also included are altering signalling through an ADP-G receptor, by
CC contacting a cell expressing the receptor with ADP-G or the ant/agonist,
CC ameliorating an ADP-G receptor associated condition, by administering a
CC therapeutic composition comprising ADP-G or the ant/agonist to an
CC individual and a composition comprising the ADP-G receptor and ADP-G.
CC ADP-G or the ant/agonist is useful for treating an ADP-G receptor
CC associated condition e.g. cardiovascular function disorder, where the
CC therapeutic composition induces vasorelaxation. The new methods are
CC useful in identifying ant/agonists and ligands of the receptor.
CC The ant/agonists are useful therapeutically for preventing or
CC ameliorating conditions associated with the receptor such as
CC cardiovascular disorders (e.g. ischemia, hypertension, hypotension,
CC angina pectoris, myocardial infarction, stroke, congestive heart
CC failure, shock, erectile dysfunction, orthostatic intolerance and
CC migraine), gastrointestinal disorders (e.g. diarrhoea, gastritis,
CC inflammatory bowel disease), immune disorders (e.g. immunodeficiency
CC disorders, autoimmune disorders, rheumatoid arthritis), infections
CC caused by bacteria, fungi, protozoa or virus, respiratory disorders
CC (e.g. asthma, pneumonia, bronchitis), kidney disorders (e.g.
CC glomerulonephritis, renal failure, lupus), hepatobiliary disorders
CC (e.g. jaundice, cirrhosis, hepatitis), endocrine disorders (e.g.
CC pituitary, thyroid or adrenal dysfunctions), musculoskeletal disorders
CC (e.g. osteoporosis, muscular dystrophies), nervous system disorders
CC (e.g. Parkinson's and Alzheimer's disease), psychotic disorders (e.g.
CC depression, anxiety, schizophrenia), pain, glycogen storage
CC diseases and disorders of body weight (e.g. as a result of cancer
CC or AIDS, acquired immunodeficiency syndrome). The gene for the ADP-G
CC receptor is located on chromosome 3 in a region associated with Usher
CC syndrome type 3 (progressive hearing loss). The ligand is useful
CC therapeutically, in detecting normal or abnormal expression of the
CC receptor in an isolated sample or in in vivo diagnostic imaging
CC procedures, and targeting specifically a diagnostic group to cells and
CC tissues that express the receptor. The present sequence encodes the
CC ADP-G receptor.
XX
SQ Sequence 1143 BP; 326 A; 257 C; 201 G; 359 T; 0 other;

Query Match 92.1%; Score 1093.8; DB 24; Length 1143;
Best Local Similarity 99.8%; Pred. No. 5.4e-280;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGGTACCAACAGAAATGCAAGCGGTGACAAATCTCCTGGCTGGGACACACAG 150
    |||||||
Db 36 AGGTACCAACAGAAATGCAAGCGGTGACAAATCTCCTGGCTGGGACACACAG 95
QY 151 TCTGTGCACGAGACTCAAAATCACCAGTCCCTCCACTGCTCTACACTGTCT 210
    |||||||
Db 96 TCTGTGCACGAGACTCAAAATCACCAGTCCCTCCACTGCTCTACACTGTCT 155
QY 211 GTTTTGTGTGACTATATCACAAATGGCTGCGATGAGATTTTCTTCAATCCGAG 270
    |||||||
Db 156 GTTTTGTGTGACTATATCACAAATGGCTGCGATGAGATTTTCTTCAATCCGAG 215
QY 271 TAAATCAAACTTATTTATTTCTTAAGAACACAGTCAATTCGTATCTGCATGATCT 330
    |||||||
Db 216 TAAATCAAACTTATTTATTTCTTAAGAACACAGTCAATTCGTATCTGCATGATCT 275
QY 331 GACTTTTCATTCAAATTTCTTAGTATGCAAACTGGGAACAGACACATGAACTTT 390
    |||||||
Db 276 GACTTTTCATTCAAATTTCTTAGTATGCAAACTGGGAACAGACACATGAACTTT 335
QY 391 TGTGTGCAATTCACCTCCGCTCATATTTTATTTTACAAATGTATACATATTTCT 450
    |||||||
Db 336 TGTGTGCAATTCACCTCCGCTCATATTTTATTTTACAAATGTATACATATTTCT 395
QY 451 GGAAGTGAATACATTCGATCGCTACAGAAAGACACAGGCACTTTAAACATCCACC 510
    |||||||
Db 396 GGAAGTGAATACATTCGATCGCTACAGAAAGACACAGGCACTTTAAACATCCACC 455
```

```
QY 511 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTCATCTGGGCAATCAATCTTACT 570
    |||
Db 456 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTCATCTGGGCAATCAATCTTACT 515
QY 571 CTTTGGCCATCATATCTGACCAAGAGCGAGCCGAGAGCAAGAAATGGAAGAATG 630
    |||
Db 516 CTTTGGCCATCATATCTGACCAAGAGCGAGCCGAGAGCAAGAAATGGAAGAATG 575
QY 631 CTTTGGCCATCAATCAAGAGTGGTGTAGTGTGATGGAATAGTAATATCATCTGTCA 690
    |||
Db 576 CTTTGGCCATCAATCAAGAGTGGTGTAGTGTGATGGAATAGTAATATCATCTGTCA 635
QY 691 AGCATATTTCTGATTAATTTCTTAATGTATGTATGTATACACTCTATTCAAAAGA 750
    |||
Db 636 AGCATATTTCTGATTAATTTCTTAATGTATGTATGTATACACTCTATTCAAAAGA 695
QY 751 ACTGTACCGGTATACATTAAGAAAGAGGGGTAGGTAAAGTCCCAAGAAAAAGGTGA 810
    |||
Db 696 ACTGTACCGGTATACATTAAGAAAGAGGGGTAGGTAAAGTCCCAAGAAAAAGGTGA 755
QY 811 CGTCAAAAGTTTCATATATCTGTATCTTATTTGTTTGTCTCTTCATTTTGC 870
    |||
Db 756 CGTCAAAAGTTTCATATATCTGTATCTTATTTGTTTGTCTCTTCATTTTGC 815
QY 871 CCGAATTCCTTACACCTGAGCCAAACCGGGAGTGTCTTGACTGCACCTGTAATAAC 930
    |||
Db 816 CCGAATTCCTTACACCTGAGCCAAACCGGGAGTGTCTTGACTGCACCTGTAATAAC 875
QY 931 TCTGTCTATGTAAGAGAGACCTGTGTGTTAACTCTTAAATGACATGCTGATCC 990
    |||
Db 876 TCTGTCTATGTAAGAGAGACCTGTGTGTTAACTCTTAAATGACATGCTGATCC 935
QY 991 GTTCATCTATTTTCTCTTGGCAAGTCTCTGAGAAATCTCTGATTAAGTCTGAAGT 1050
    |||
Db 936 GTTCATCTATTTTCTCTTGGCAAGTCTCTGAGAAATCTCTGATTAAGTCTGAAGT 995
QY 1051 CCCCAATTCGCAACATCTCTGTCCAGAGCAATAGAAAAAAGAACAGAGTGTGTGA 1110
    |||
Db 996 CCCCAATTCGCAACATCTCTGTCCAGAGCAATAGAAAAAAGAACAGAGTGTGTGA 1055
QY 1111 CCCAATGAGAGACTCAATGTAACAATAATTAAGTAATTTTCAATCTCTTGT 1170
    |||
Db 1056 CCCAATGAGAGACTCAATGTAACAATAATTAAGTAATTTTCAATCTCTTGT 1115
QY 1171 GTTCAGACTCGTTAA 1187
    |||
Db 1116 GTTCAGACTCGTTAA 1132

RESULT 11
AA168802
ID AA168802 standard; DNA; 2245 BP.
XX
AC AA168802;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human PAFR3 DNA.
XX
KW PAFR-3; platelet activating factor receptor; human; chromosome 3;
KW thrombocyte activation; hypotension; vascular permeability;
KW bronchoconstriction; transgenic animal; knockout animal; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 467..1195
FT /tag= a
FT /product= "PAFR3"
XX
PN DE10020073-A1.
XX
PD 25-OCT-2001.
```

```
XX 22-APR-2000; 2000DE-1020073.
PF 22-APR-2000; 2000DE-1020073.
XX
PR 22-APR-2000; 2000DE-1020073.
XX
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.
XX
PI Brues M, Boenisch H;
XX
PI WPI: 2002-027296/04.
XX
DR P-PDB; AAG80236.
XX
PT New human platelet-activating factor (PAF) receptor-3 gene, useful for
PT diagnosis and treatment of PAF-related diseases
XX
PS Disclosure: Page 3-4; 6pp; German.
XX
CC This invention describes a novel human platelet-activating factor (PAF)
CC receptor-3 (PAFR-3) gene (I). The protein encoded by (I) probably
CC modulates the activity of PAF, which is involved in many
CC (patho)physiological processes, e.g., thrombocyte activation,
CC hypotension, increased vascular permeability, bronchoconstriction etc.
CC (I), and derived (anti)sense oligonucleotides, are useful in the
CC treatment and diagnosis of (I)-related diseases; for producing
CC transgenic/knockout animals, and for recombinant expression of the
CC protein (II) that it encodes. (II) is useful in ligand-binding studies
CC and screening assays, also for treatment and diagnosis of (II)-related
CC diseases. This sequence encodes the human leucocyte platelet-activating
CC factor receptor (PAFR-3) protein described in the invention.
XX
SQ Sequence 2245 BP; 676 A; 466 C; 397 G; 706 T; 0 other;
Query Match 92.1%; Score 1093.8; DB 24; Length 2245;
Best Local Similarity 99.8%; Pred No. 7e-280;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 91 AGGTAACCAAGAAATGCAAGCCGTCGACATCTCACCTGCGCTGGGAACACAG 150
    |||
Db 451 AGGTAACCAAGAAATGCAAGCCGTCGACATCTCACCTGCGCTGGGAACACAG 150
QY 151 TCTGTGACACAGAGACTACAAAATCAACAGGTCCTTCCACCTGCTACTACTGCT 210
    |||
Db 511 TCTGTGACACAGAGACTACAAAATCAACAGGTCCTTCCACCTGCTACTACTGCT 210
QY 211 GTTTTGTGACATATCAACAATGGCTGGGATGAGATTTTCTTCAATCCGAG 270
    |||
Db 571 GTTTTGTGACATATCAACAATGGCTGGGATGAGATTTTCTTCAATCCGAG 270
QY 271 TAAATCAACTTATTTATTTTCTTAAGAACACAGTCATTTCTGATCTCTCATGATCT 330
    |||
Db 631 TAAATCAACTTATTTATTTTCTTAAGAACACAGTCATTTCTGATCTCTCATGATCT 330
QY 331 GACTTTTCATTCAAAATTTCTAGTATGCCAAATGCGGAACAGGACCACTGAACTTT 390
    |||
Db 691 GACTTTTCATTCAAAATTTCTAGTATGCCAAATGCGGAACAGGACCACTGAACTTT 390
QY 391 TGTGTGACAGTACCTCGCTCATATTTTATTTACAAATGATATAGTATTCATTCCT 450
    |||
Db 751 TGTGTGACAGTACCTCGCTCATATTTTATTTACAAATGATATAGTATTCATTCCT 450
QY 451 GGGAGTGAATACATGATGATGCTACAGAAAGACACACAGCCATTTAAACATCCAAACC 510
    |||
Db 811 GGGAGTGAATACATGATGATGCTACAGAAAGACACACAGCCATTTAAACATCCAAACC 510
QY 511 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTCATCTGGGCAATCAATCTTACT 570
    |||
Db 871 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTCATCTGGGCAATCAATCTTACT 570
QY 571 CTTTGGCCATCATATCTGACCAAGAGCGAGCCGAGAGCAAGAAATGGAAGAATG 630
    |||
Db 931 CTTTGGCCATCATATCTGACCAAGAGCGAGCCGAGAGCAAGAAATGGAAGAATG 630
```

Query Match	Best Local Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
107	ATCGAAGCCGTCGACATCTCAGCTTCGCGCTGGGGAACCAACAGTCTGTGCACGAGAC	166							
1	ATGGAAGCCGTCGACAAATCTCAGCTTCGCGCTGGGGAACCAACAGTCTGTGCACGAGAC	60							
167	TACAAATCACCAGGTCCTCTCCACATGCTGACATGTCGTCTTTTGTGGAGCTT	226							
61	TACAAATCACCAGGTCCTCTCTCCACATGCTGACATGTCGTCTTTTGTGGAGCTT	120							
227	ATCACAATGCGCTGCGCATGAGATTTCTTTCATTAATCCGGAGTAATCAAACTTTATT	286							
121	ATCACAATGCGCTGCGCATGAGATTTCTTTCATTAATCCGGAGTAATCAAACTTTATT	180							
287	ATTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTCTGACTTTTCATTCAAA	346							
181	ATTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTCTGACTTTTCATTCAAA	240							
347	ATTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTCTGACTTTTCATTCAAA	406							
241	ATTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTCTGACTTTTCATTCAAA	300							

```

QY 407 TCGTCATATTTATTTACAAATGATATGATATTTTCATTCCTGGGACGTATACATC 466
    |||
Db 301 TCGTCATATTTATTTATTTACAAATGATATGATATTTTCATTCCTGGGACGTATACATC 360
QY 467 GATGCGACCAAGAGACACACGACCATTTAAACATCCAAACCCCAAAATCTCTGGG 526
    |||
Db 361 GATGCGACCAAGAGACACACGACCATTTAAACATCCAAACCCCAAAATCTCTGGG 420
QY 527 GCTAAGATTCCTCTGTTGTCATCTGGGCAATCATCTCTTCTCTGGGCAATG 586
    |||
Db 421 GCTAAGATTCCTCTGTTGTCATCTGGGCAATCATCTCTTCTCTGGGCAATG 480
QY 587 ATTCTGACCAACAGCAGCCGAGAGACAAAGATGGAAGAAATCTCTTCTTAATCA 646
    |||
Db 481 ATTCTGACCAACAGCAGCCGAGAGACAAAGATGGAAGAAATCTCTTCTTAATCA 540
QY 647 GAGTTCGCTAGTCTGCGCATGAAATAGTAATACATCTCTCAAGTCATTTCTGAT 706
    |||
Db 541 GAGTTCGCTAGTCTGCGCATGAAATAGTAATACATCTCTCAAGTCATTTCTGAT 600
QY 707 AATTTCTTAATTTGTAATGTAATGTAATACATCTCAAAAGAACTGTACCGGTATAC 766
    |||
Db 601 AATTTCTTAATTTGTAATGTAATGTAATACATCTCAAAAGAACTGTACCGGTATAC 660
QY 767 GTAAGACGAGGGGTGAGGTAAGTCCCAAGAAAGGTAAGGTAAGTCCCAAGTTCAT 826
    |||
Db 661 GTAAGACGAGGGGTGAGGTAAGTCCCAAGAAAGGTAAGGTAAGTCCCAAGTTCAT 720
QY 827 ATCATTCGCTATTCCTTAATTTGTTGTTCTCTTCCATTTGGCCCAATTCCTTACAC 886
    |||
Db 721 ATCATTCGCTATTCCTTAATTTGTTGTTCTCTTCCATTTGGCCCAATTCCTTACAC 780
QY 887 CTGAGCCAAACCCGGGATGCTTTGACTGCACTGCTGAATAATCTCTGTTCTATGTA 946
    |||
Db 781 CTGAGCCAAACCCGGGATGCTTTGACTGCACTGCTGAATAATCTCTGTTCTATGTA 840
QY 947 GAGAGCACTGCTGTTGTAATCTCTTAATGCAATGCTGGATCCGTCATCTATTTTTC 1006
    |||
Db 841 GAGAGCACTGCTGTTGTAATCTCTTAATGCAATGCTGGATCCGTCATCTATTTTTC 900
QY 1007 CTTTGCAGTCTCTCAGAAATCTCTGATATGATGCTGAAGTCCCAATTCGCAACA 1066
    |||
Db 901 CTTTGCAGTCTCTCAGAAATCTCTGATATGATGCTGAAGTCCCAATTCGCAACA 960
QY 1067 TCTCTGTCACAGACATAGGAAAAAAGAACAGATGCTGTCACCAATGAGAGACT 1126
    |||
Db 961 TCTCTGTCACAGACATAGGAAAAAAGAACAGATGCTGTCACCAATGAGAGACT 1020
QY 1127 CCAATGTAA 1135
    |||
Db 1021 CCAATGTAA 1029

```

RESULT 13
 ID AAA46034 standard; cDNA; 1029 BP.

XX AAA46034;
 DT 22-AUG-2000 (first entry)

XX Human G protein coupled receptor hcHNB encoding cDNA seq ID NO:33.

KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW Identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant; ss.

OS Homo sapiens.

PN WO200022131-A2.

PD 20-APR-2000.

```

PF 13-OCT-1999; 99WO-US24065.
XX
PR 13-OCT-1998; 98US-0170496.
PR 12-NOV-1998; 98US-0108029.
PR 20-NOV-1998; 98US-0109213.
PR 27-NOV-1998; 98US-0110060.
PR 16-FEB-1999; 99US-0120416.
PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-MAY-1999; 99US-0137567.
PR 30-JUN-1999; 99US-0141448.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156533.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.

```

(AREN-) ARENA PHARM INC.

Behan DP, Lehmann-Brylinsma K, Chalmers DT, Chen R, Dang HT;
 Gore M, Llaw CW, Lin I, Lowitz K, White C;

WPI; 2000-317986/27.

P-PSDB; AAB02840.

Non-endogenous, human G protein-coupled receptors for screening
 receptor, inverse or partial agonists useful as therapeutic agents

Example 1; Page 111-112; 187pp; English.

The present invention describes transmembrane receptors, preferably
 human G protein coupled receptors (GPCR), for which the endogenous
 ligand is unknown (orphan GPCR receptors). More specifically the present
 invention relates to non-endogenous, constitutively activated versions
 of a human GPCR. These non-endogenous human GPCRs can be useful for
 the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 the exemplification of the present invention.

Sequence 1029 BP; 285 A; 233 C; 186 G; 325 T; 0 other;

Query Match 86.7%; Score 1029; DB 21; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 7.9e-263;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 107 ATGCAAGCGCTGACAAATCTACCTGCGCTGGGAACACAGTCTGTCAACAGAC 166
    |||
Db 1 ATGCAAGCGCTGACAAATCTACCTGCGCTGGGAACACAGTCTGTCAACAGAC 60
QY 167 TACAAATACCCAGGCTCTCTCCACCTGCTACACCTGCTCTTCTTTTGTGGACT 226
    |||
Db 61 TACAAATACCCAGGCTCTCTCCACCTGCTACACCTGCTCTTCTTTTGTGGACT 120
QY 227 ATCACAATGCGCGGAGAGAGATTTCTTCAAAATCCGAGTAATCAACTTAT 286
    |||
Db 121 ATCACAATGCGCGGAGAGAGATTTCTTCAAAATCCGAGTAATCAACTTAT 180
QY 287 ATTTTCTTAAGAACACAGTCTGATCTCTCATGATCTGACTTTTCCATTCAA 346
    |||
Db 181 ATTTTCTTAAGAACACAGTCTGATCTCTCATGATCTGACTTTTCCATTCAA 240
QY 347 ATTCTAGTATGCCAAATGAGAACAGACACTGAGAACTTTTGTGTCAAGTTACC 406

```


|||||
Db 541 GAGTTCGCTAGCTGCGCATGAAATGTAATTAATACATCTGTCAAGCATTTTCTGCAAT 600
QY 707 AATTTCTTAATGTTATGTTATGTTATATACATCATATACAAAGAACTGTACCGGTACATAC 766
Db 601 AATTTCTTAATGTTATGTTATGTTATATACATCATATACAAAGAACTGTACCGGTACATAC 660
QY 767 GTAAAGAACGAGGGGTGTAGTAAAGTCCCAAGGAAAAAGTGAACGTCAAAAGTTTCATT 826
Db 661 GTAAAGAACGAGGGGTGTAGTAAAGTCCCAAGGAAAAAGTGAACGTCAAAAGTTTCATT 720
QY 827 ATCATTCCTGATTTCTTATTTGTTGTTGTTCTTCCATTTTGCCGGAATTCCTTACACC 886
Db 721 ATCATTCCTGATTTCTTATTTGTTGTTGTTCTTCCATTTTGCCGGAATTCCTTACACC 780
QY 887 CTGAGCCAAACCCGGAGTCTTGTGACTGCATGCTGAAATATCTGTTCTATGTGAAA 946
Db 781 CTGAGCCAAACCCGGAGTCTTGTGACTGCATGCTGAAATATCTGTTCTATGTGAAA 840
QY 947 GAGAGCACTCTGTGTTAACTTCTTAATGATGCTGATCCGTTCAATCTATTTTTC 1006
Db 841 GAGAGCACTCTGTGTTAACTTCTTAATGATGCTGATCCGTTCAATCTATTTTTC 900
QY 1007 CTTTGCAAGTCTTCAAGAAATCTTGATAGATGCTGAAGTGGCCCAATTCGACACA 1066
Db 901 CTTTGCAAGTCTTCAAGAAATCTTGATAGATGCTGAAGTGGCCCAATTCGACACA 960
QY 1067 TCTCTGTCCAGAGCAATAGGAAAAAGAACAGAGTGGTGTGACCCCAATGAGAGACT 1126
Db 961 TCTCTGTCCAGAGCAATAGGAAAAAGAACAGAGTGGTGTGACCCCAATGAGAGACT 1020
QY 1127 CCAATGTAA 1135
Db 1021 CCAATGTAA 1029

RESULT 15

ABA98535

ID ABA98535 standard; cDNA; 1029 BP.

AC ABA98535;

DT 25-APR-2002 (first entry)

DE Human G protein-coupled receptor, IGPcR17, coding sequence.

XX Human; G protein-coupled receptor; IGPcR17; gene; analgesic; neuroleptic;

KM tranquiliser; antiparkinsonian; neuroprotective; nootropic;

KM anticonvulsant; metabolic; anorectic; anabolic; antiinflammatory;

KM antiarthritic; osteopathic; antiasthmatic; antiallergic; antidiabetic;

KM immunosuppressive; gene therapy; psychiatric disorder; schizophrenia;

KM central nervous system disorder; movement dysfunction; schizophrenia;

KM multiple sclerosis; Alzheimer's disease; kidney disease; obesity;

KM gastrointestinal disorder; osteoporosis; infection;

KM gynecological disorder; ss.

OS Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

PI Wattler F, Wattler S, Trommler P, Nehls WC;
XX MPI; 2002-14080/18.
DR P-PSDB; AAM48353.
XX
PT New human or mouse G protein-coupled receptor protein, IGPcR17, useful
PT for diagnosis, prevention, amelioration or treatment of central nervous
PT system disorders such as Tourette's syndrome, Parkinson's disease and
PT pain
PS Claim 1; Fig 1; 71pp; English.

CC The present sequence is the coding sequence for human G protein-coupled
CC receptor (IGPcR) protein, IGPcR17. The coding sequence for IGPcR17 is
CC useful in gene therapy for prevention, amelioration or treatment of
CC diseases characterised by aberrant expression or activity of IGPcR17,
CC where the disease is a psychiatric or central nervous system (CNS)
CC disorder associated with signal processing in CNS such as learning and
CC memory disorders, movement dysfunctions, tics, tremor, Tourette's
CC syndrome, Parkinson's disease, Huntington's disease, dyskinesias,
CC dystonia, pain and spasms. In addition, IGPcR17 and its coding sequence
CC are useful in diagnosis, prevention, amelioration or treatment of
CC diseases associated with signal processing in CNS, schizophrenia,
CC episodic paroxysmal anxiety (EPA) disorders such as obsessive compulsive
CC disorder (COD), multiple sclerosis, Alzheimer's disease/dementia,
CC anorexia, kidney diseases such as renal failure, obesity,
CC gastrointestinal disorders such as irritable bowel syndrome (IBS),
CC diarrhoea, motility disorders and conditions of delayed gastric emptying,
CC osteoporosis, infections such as bacterial, fungal, protozoal and viral
CC infections, asthma, allergy, arthritis, sepsis and gynecological
CC disorders.

XX Sequence 1029 BP; 285 A; 234 C; 186 G; 324 T; 0 other;

Query Match 86.6%; Score 1027.4; DB 24; Length 1029;

Best Local Similarity 99.9%; Pred. No. 2,1e-262;

Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 107 ATGCAAGCCGTCGACAAATCTACCTCTGCGCTGGGAACACAGTGTGACACAGAGAC 166
Db 1 ATGCAAGCCGTCGACAAATCTACCTCTGCGCTGGGAACACAGTGTGACACAGAGAC 60
QY 167 TACAAAATGACCCAGGTCCTTCCACCTGCTTACACTGTCTGTTTGTGGACTT 226
Db 61 TACAAAATGACCCAGGTCCTTCCACCTGCTTACACTGTCTGTTTGTGGACTT 120
QY 227 ATGCAAAATGCGCTGGGAGAGAGATTTCTTCAAAATCGGAGTAAATCAATTAAT 286
Db 121 ATGCAAAATGCGCTGGGAGAGAGATTTCTTCAAAATCGGAGTAAATCAATTAAT 180
QY 287 ATTTTCTTAAAGAACACAGTCAATTTCTGATCTTCTGATGATCTGACTTTTCCATTC 346
Db 181 ATTTTCTTAAAGAACACAGTCAATTTCTGATCTTCTGATGATCTGACTTTTCCATTC 240
QY 347 ATTTCTTATGATGCGCAATGCGGAGACAGACCTGAGAACTTTGTGTCAAGTTACC 406
Db 241 ATTTCTTATGATGCGCAATGCGGAGACAGACCTGAGAACTTTGTGTCAAGTTACC 300
QY 407 TCCGTCATATTTTATTTTCAATATGATATATGATATTTTCAATTCGCGGAGATTAAT 466
Db 301 TCCGTCATATTTTATTTTCAATATGATATATGATATTTTCAATTCGCGGAGATTAAT 360
QY 467 GATGCTACAGAGAGACAGCCAGCTTTTAAACATCAACCCCAAAATCTCTTGGGG 526
Db 361 GATGCTACAGAGAGACAGCCAGCTTTTAAACATCAACCCCAAAATCTCTTGGGG 420
QY 527 GCTAAGATTTCTCTGTGTCATCTGGGATTTATGTTCTTACTCTTGTGCTTAACATG 586
Db 421 GCTAAGATTTCTCTGTGTCATCTGGGATTTATGTTCTTACTCTTGTGCTTAACATG 480
QY 587 ATTCTGACCAAGAGGAGCGAGAGACAAAGATGTAAGAAATGCTTTTCTTAATCA 646
Db 481 ATTCTGACCAAGAGGAGCGAGAGACAAAGATGTAAGAAATGCTTTTCTTAATCA 540

```
QY 647 GAGTCGGTCTAGTCTGCGATGAATAAGTAATACATCTGCAAGTCATTTTCGATT 706
    |||||||
Db 541 GAGTTCGGTCTAGTCTGCGATGAATAAGTAATACATCTGCAAGTCATTTTCGATT 600
    |||||||
QY 707 AATTCTTAATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 766
    |||||||
Db 601 AATTCTTAATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 660
    |||||||
QY 767 GTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 826
    |||||||
Db 661 GTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
    |||||||
QY 827 ATCATTCGCTGATTCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 886
    |||||||
Db 721 ATCATTCGCTGATTCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 780
    |||||||
QY 887 CTGAGCCAAACCCGGGATGCTTTGACTGCACTGCTGAAATACTGTTCTATGTGAAA 946
    |||||||
Db 781 CTGAGCCAAACCCGGGATGCTTTGACTGCACTGCTGAAATACTGTTCTATGTGAAA 840
    |||||||
QY 947 GAGAGCACTCTGAGTTAACTTCTTAATGATGCCGATCCGTTGATCTATTTTTC 1006
    |||||||
Db 841 GAGAGCACTCTGAGTTAACTTCTTAATGATGCCGATCCGTTGATCTATTTTTC 900
    |||||||
QY 1007 CTTTGCAAGTCTTCAGAAATTCCTGATAAGTGAAGTGCCCAATTCGCAACA 1066
    |||||||
Db 901 CTTTGCAAGTCTTCAGAAATTCCTGATAAGTGAAGTGCCCAATTCGCAACA 960
    |||||||
QY 1067 TCTCTGTCCCGAGACATAGGAAAAAGACAGAGTGGTGAACCCAAATGAAGAGACT 1126
    |||||||
Db 961 TCTCTGTCCCGAGACATAGGAAAAAGACAGAGTGGTGAACCCAAATGAAGAGACT 1020
    |||||||
QY 1127 CCAATGTAA 1135
    |||||||
Db 1021 CCAATGTAA 1029
```

Search completed: December 11, 2002, 06:08:42
Job time : 331 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 03:17:19 ; Search time 75 Seconds

(Without alignments)
4853.671 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187

Sequence: 1 ggtaccggggatcctctag.....tgtgtcagaactcgtaaa 1187

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1160	97.7	2247	3 US-08-852-824-1	Sequence 1, Appl
2	305.6	25.7	1124	4 US-09-221-456-1	Sequence 1, Appl
3	305.6	25.7	1124	4 US-09-358-740-1	Sequence 1, Appl
4	302.4	25.5	1488	2 US-08-812-871-2	Sequence 2, Appl
5	295.4	24.9	1474	2 US-08-467-948A-5	Sequence 5, Appl
6	295.4	24.9	1474	3 US-08-467-947A-5	Sequence 5, Appl
7	279	23.5	302	4 US-08-905-223-94	Sequence 94, Appl
8	250.8	21.1	1027	4 US-09-303-824A-1	Sequence 1, Appl
9	207.4	17.5	1444	3 US-08-968-876-4	Sequence 4, Appl
10	129	10.9	1243	1 US-08-702-344-27	Sequence 27, Appl
11	77.2	6.5	1143	2 US-08-467-125-1	Sequence 1, Appl
12	77.2	6.5	1143	2 US-08-911-320A-1	Sequence 1, Appl
13	77.2	6.5	1143	3 US-09-217-101-1	Sequence 1, Appl
14	71.4	6.4	414	3 US-08-513-874B-353	Sequence 353, App
15	71.4	6.0	357	3 US-08-513-974B-51	Sequence 51, Appl
16	63.4	5.3	1102	2 US-08-742-440A-5	Sequence 5, Appl
17	63.4	5.3	1224	2 US-08-742-440A-4	Sequence 4, Appl
18	62	5.2	1572	1 US-08-041-219A-5	Sequence 5, Appl
19	62	5.2	1572	1 US-08-417-122-5	Sequence 5, Appl
20	61.8	5.1	1080	4 US-09-050-159-133	Sequence 133, App
21	60.2	5.1	1080	3 US-08-602-809-3	Sequence 3, Appl
22	60.2	5.1	1080	3 PCT-US95-16472-3	Sequence 3, Appl
23	56.8	4.8	1155	3 US-09-053-866-3	Sequence 3, Appl
24	56.8	4.8	1155	4 US-09-479-130-3	Sequence 3, Appl
25	52.6	4.4	1224	2 US-08-742-440A-1	Sequence 1, Appl
26	52.6	4.4	1235	1 US-08-696-770-1	Sequence 1, Appl
27	52.6	4.4	1235	2 US-09-015-557-1	Sequence 1, Appl

28	50	4.2	1083	4 US-09-116-498-7	Sequence 7, Appl
29	48.4	4.1	1083	4 US-09-116-498-11	Sequence 11, Appl
30	47.6	4.0	1301	2 US-08-467-948A-7	Sequence 7, Appl
31	47.6	4.0	1301	3 US-08-467-947A-7	Sequence 7, Appl
32	47.4	4.0	1610	5 US-08-889-108-7	Sequence 7, Appl
33	47.4	4.0	1610	5 PCT-US94-10358-7	Sequence 7, Appl
34	47.4	4.0	2160	4 US-08-188-275A-1	Sequence 1, Appl
35	47.4	4.0	2162	4 US-09-351-198-1	Sequence 1, Appl
36	47.4	4.0	2162	4 US-09-113-426-1	Sequence 1, Appl
37	45.8	3.9	1317	1 US-08-153-848-45	Sequence 45, Appl
38	45.8	3.9	1317	3 US-09-289-843A-45	Sequence 45, Appl
39	45.8	3.9	1317	4 US-09-088-337B-45	Sequence 45, Appl
40	45.8	3.9	1317	5 PCT-US93-11153-45	Sequence 45, Appl
41	45.8	3.9	1664	4 US-09-582-224A-5	Sequence 5, Appl
42	45.8	3.9	1679	4 US-09-517-605-14	Sequence 14, Appl
43	45.8	3.9	1737	1 US-08-202-056-4	Sequence 4, Appl
44	45.8	3.9	1737	1 US-08-076-093A-3	Sequence 3, Appl
45	45.8	3.9	1737	1 US-08-701-263-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1									
US-08-852-824-1									
; Sequence 1, Application US/08852824C									
; Patent NO. 6060272									
; GENERAL INFORMATION:									
; APPLICANT: Li et al.									
; TITLE OF INVENTION: Human G-Protein Coupled Receptors									
; FILE REFERENCE: 1488.122000									
; CURRENT APPLICATION NUMBER: US/08/852,824C									
; CURRENT FILING DATE: 1997-05-04									
; NUMBER OF SEQ. ID NOS: 18									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 1									
; LENGTH: 2247									
; TYPE: DNA									
; ORGANISM: genomic									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (226)..(1251)									
US-08-852-824-1									
Query Match									
Best Local Similarity 97.7%; Score 1160; DB 3; Length 2247;									
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	28	CTGCAGCTGCATTA	CTACTACTGATGATCAATCAACCCCTCAGAAATCAACAGTT	87					
DB	147	CTGCAGCTGCATTA	CTACTACTGATGATCAATCAACCCCTCAGAAATCAACAGTT	206					
QY	88	ATCAGGTACCAACAA	GAATGCAAGCCGTGCAATCTCACCTTGGCGCTGGGAACAC	147					
DB	207	ATCAGGTACCAACAA	GAATGCAAGCCGTGCAATCTCACCTTGGCGCTGGGAACAC	266					
QY	148	CAGTCTGACACAG	AGACTCAAAATCACCAGGCTCTCCACAGCTCTACACTGT	207					
DB	267	CAGTCTGACACAG	AGACTCAAAATCACCAGGCTCTCTCCACAGCTCTACACTGT	326					
QY	208	CCTGTTTTTGTG	ACTTATCACAATGGCTGGCGATGAGGATTTCTTCAATCCG	267					
DB	327	CCTGTTTTTGTG	ACTTATCACAATGGCTGGCGATGAGGATTTCTTCAATCCG	386					
QY	268	GAGTAAATCAAC	TTTATTTCTTAAACACATCTTGTGATCTTTCATGAT	327					
DB	387	GAGTAAATCAAC	TTTATTTCTTAAACACATCTTGTGATCTTTCATGAT	446					
QY	328	TCTGACTTTTCAT	TCAAAATTTCTAGTATGCCAAATGGGAACAGACCACTAGAAC	387					
DB	447	TCTGACTTTTCAT	TCAAAATTTCTAGTATGCCAAATGGGAACAGACCACTAGAAC	506					
QY	388	TTTGTGTGTCAG	ATTTACCTCGTCATATTTTATTTTCAAAATGATATACATATTTCAT	447					

|||||
Db 507 TTTTGTGTGTCAGTACCTCCCTCATATTTTTCACATGTATATCGTATTTTCATT 566
Qy 448 CTTGGGACTGATATCATATCGTACACGAGAACCCAGCCATTTTAAACATCCAA 507
Db 567 CTGTGGACTGATATCATATCGTACACGAGAACCCAGCCATTTTAAACATCCAA 626
Qy 508 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTCTGTGTATCTGGGCAATTCATCTT 567
Db 627 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTCTGTGTATCTGGGCAATTCATCTT 686
Qy 568 ACTCTCTTTCCTAATCATATGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAA 627
Db 687 ACTCTCTTTCCTAATCATATGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAA 746
Qy 628 ATGCTCTTTCCTAATCATATGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAA 687
Db 747 ATGCTCTTTCCTAATCATATGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAA 806
Qy 688 TCAAGCATTTTCTGATTAATTTCTTAATGTATGTATGTATGTATGTATGTATGTAT 747
Db 807 TCAAGCATTTTCTGATTAATTTCTTAATGTATGTATGTATGTATGTATGTATGTAT 866
Qy 748 AGAAGTATACCGGTATACGTAAAGAGAGAGGAGTGTAGTAAGTCCCGAGAAAAAGT 807
Db 867 AGAAGTATACCGGTATACGTAAAGAGAGAGGAGTGTAGTAAGTCCCGAGAAAAAGT 926
Qy 808 GAAAGTCAAAATTTTCAATATATATATATATATATATATATATATATATATATAT 867
Db 927 GAAAGTCAAAATTTTCAATATATATATATATATATATATATATATATATATATAT 986
Qy 868 TGCCCGAATTCCTTACACCCGAGCCAAACCCGAGATGCTTGTACATGCTCTGAAAA 927
Db 987 TGCCCGAATTCCTTACACCCGAGCCAAACCCGAGATGCTTGTACATGCTCTGAAAA 1046
Qy 928 TACTGTCTATGTAAAGAGAGACACTCTGTGTAACTTCACTTAAATGCATGCCGTA 987
Db 1047 TACTGTCTATGTAAAGAGAGACACTCTGTGTAACTTCACTTAAATGCATGCCGTA 1106
Qy 988 TCCGTTTCAATATTTTCTTTCGCAAGTCTTTCAGAAATTCCTGTAAATATGCTGAA 1047
Db 1107 TCCGTTTCAATATTTTCTTTCGCAAGTCTTTCAGAAATTCCTGTAAATATGCTGAA 1166
Qy 1048 GTGCCCCAATTCGCAACATCTGTGCCAGAGCAATAGAAAAAGAAAGAGATGTGG 1107
Db 1167 GTGCCCCAATTCGCAACATCTGTGCCAGAGCAATAGAAAAAGAAAGAGATGTGG 1226
Qy 1108 TGACCCAAATGAGAGAGACTCCATGTAAACAATTAATTAATTAATTAATTAAT 1167
Db 1227 TGACCCAAATGAGAGAGACTCCATGTAAACAATTAATTAATTAATTAATTAAT 1286
Qy 1168 TGTGTTCAGAGACTGTTAAA 1187
Db 1287 TGTGTTCAGAGACTGTTAAA 1306

RESULT 2
US-09-221-456-1
Sequence 1, Application US/09221456
Patent No. 6162899
GENERAL INFORMATION:
APPLICANT: SATHE, GANESH
APPLICANT: HALSEY, WENDY
APPLICANT: MOIR, ALISON
APPLICANT: CHAMBERS, JON
APPLICANT: SEKERES, PHILIP
TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS OF THE HNEA81 RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge

STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,456
FILING DATE: 28-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,975
FILING DATE: 23-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70318-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-221-456-1

Query Match 25.7%; Score 305.6; DB 4; Length 1124;
Best Local Similarity 57.7%; Pred. No. 8.4e-80;
Matches 545; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

Qy 152 CTGTGACCCAGAGACATCAAAATACACCCAGTCTCTCCATCTGTACATGTCGCG 211
Db 137 CGGTGCCCCAGAGACATCGAGATGATGATGATGATGATGATGATGATGATGATG 196
Qy 212 TTTTGTGTGAGTATATCAAAATGCGCTGCGATGAGATTTTCTTCAATCCGAGT 271
Db 197 TTCTTACCGGATCTCTGCGATGATGATGATGATGATGATGATGATGATGATGAT 256
Qy 272 AAATCAAACTTATATATTTTCTTAAGAACACAGTATTTCTGATCTGTCAATTCG 331
Db 257 TCCCTCACCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 316
Qy 332 ACTTTCACATCAAAATTTCTAGTATGATGATGATGATGATGATGATGATGATGAT 391
Db 317 ATGCTTCCTTCAAAATCTCTCTGACTCACACCTGACACCTGACCTGACAGCTTTT 376
Qy 392 GTGTGTCAAGTATCTCCGTCATATTTTATTTTCAATGATATATATATATATATAT 451
Db 377 GTGTGTCAAGTATCTCCGTCATATTTTATTTTGAAGCATGATGATGATGATGATG 436
Qy 452 GGAAGTATATATATATATATATATATATATATATATATATATATATATATATAT 511
Db 437 GGGCTCATAGCCCTTTCAGAGATTCCTCAAGATCATCAGACCTTTGAGAAATATTT 496
Qy 512 AAAAATCTTGGGGGCTAAGATTCCTCTGTGTGATGATGATGATGATGATGATGAT 571
Db 497 AAAAATCTTGGGGGCTAAGATTCCTCTGTGTGATGATGATGATGATGATGATGAT 556
Qy 572 TCTTTCCTAATCATATGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAAATGC 631
Db 557 TCCCTGCAAAATCATATCTTGTGACCAACAGCAGCCAGAGACAAAGATGTAAGAA 616
Qy 632 TCTTTCCTAATCATATGTGAGTGTGATGATGATGATGATGATGATGATGATGAT 691
Db 617 GCTTTCCTAAGAGGCTCTGGGGCTGAAGATGATGATGATGATGATGATGATGATG 676
Qy 692 GTCATTTTCTGATTAATTTCTTAATTTGATGATGATGATGATGATGATGATGATG 751


```

Db 336 CTTTGTGTGCTGTTTCTTGGTGATATTTATGAGACCATGTATGTGGCATGTGC 395
Qy 447 TCTGGAGTATACATATGATCGCTACCAAGACACCCAGGCCATTTAAACATCA 506
Db 396 TGTTAGGGCATACGCTTGTACAGATTCCTCAAGATCATACACCTTTGAGAAATATT 455
Qy 507 ACCCAAAATCTCTGGGGCTAAGATTCCTGTGTGATCTGGGCACTTATGTCT 566
Db 456 TTCTAAAAAAGCTGTTGGGAAAAAGGCTCAATCTTCATCTGGTCTTTGGTCT 515
Qy 567 TACTCTCTGCTTAACATGATTCGTACCAACAGAGCCGAGACAAAGATGTGAGA 626
Db 516 TCATCTCCCGCAAAATATATCTGTAGCAACAGAGAACACACATCTCTGTGAAA 575
Qy 627 AATGCTCTTCTTAATACAGATTCGGTCTAGTGGCAATGTAATATTAATCTCT 686
Db 576 AGTGTCTCTTAAAGGGGCTCTGTGGCTGAATGGCTCAAAATGGTCAATAT 635
Qy 687 GTCAAGTCATTTCTGGATTAATTTCTTAATGTATATGTATGATACATCA 746
Db 636 GCCAGTTATTTCTGACTGTTTATCTTAATGCTTGTGTTTATGTGTTATGCAA 695
Qy 747 AAGAGTGTACCGGTATACGTAGAAGAGGAGGTAGGTAAAGTCCAGGAAAAAG 806
Db 696 AA-AGATATATGATTTCTATAGAAAGTCCAAAGTAAAGACAAAAAACAAAAAGC 754
Qy 807 TGAACGTCAAGTTTTCATTAATCATGTCGTATTCCTTAATTTGTTTCTTCCATT 866
Db 755 TGGAAAGCAAGATTTGTTGTGCTGCTCTCTTGTGTTGTTGCTCCATTTTCAT 814
Qy 867 TTTCCCGAATTCCTTACACCTGAGCCAAACCCGGATGCTTGTAGTGCATGCTGAA 926
Db 815 TCCCGAGAGTTCATATACACAGTCAACCAACAAATAGACTGTAGTGAAGTCAAA 874
Qy 927 ATACTCTGTTCTTGAAGAGACACTCTGTGTTAACTTCTTAATGATGCTGCTG 986
Db 875 ATCACTGTTATTTGTAAGAAACAACTCTCTTTTGGAGACATAATTTGATG 934
Qy 987 ATCCGTTCACTATTTTCTCTTCTTGAAGTCTTCAAGAAATCTCTTGAATGATGCTGA 1046
Db 935 ATCCCTTAATATACATATCTTATGTAAAAAATTCACAGAAAAAGTCACTATGATGCAAG 994
Qy 1047 AGTGGCCCAATCTGCAACATCTCTGCCAGCAATAGGAAAAAGA 1095
Db 995 GGAAGAAAGCACACGATCAAGCAAGAAAAATCATAGCAGTACAGACAGA 1043

RESULT 7
US-08-905-223-94
; Sequence 94, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: s19_deptide
; LOCATION: 99..236
; IDENTIFICATION METHOD: Von Heljne matrix
; OTHER INFORMATION: score 5.4
; OTHER INFORMATION: seq VLEFVGLTNGLA/MR
US-08-905-223-94

Query Match 23.5%; Score 279; DB 4; Length 302;
Best Local Similarity 99.3%; Pred. No. 2,6e-72;
Matches 279; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 28 CTCGAGGCTGCATTAATCTACTACTGATATCATTTCAAAACCTTCAGAAATCAAGTT 87
Db 20 CTCGAGGCTGCATTAATCTACTACTGATATCATTTCAAAACCTTCAGAAATCAAGTT 79
Qy 88 ATCAGGTAACCAAGAAATGCAAGCCGTCGACATCTACCTGCGCTGGAAAC 147
Db 80 ATCAGGTAACCAAGAAATGCAAGCCGTCGACATCTACTCTGCGCTGGAAAC 139
Qy 148 CAGTCTGTGACACGAGACTACAAATCACCCAGTCTCTCCACTGCTTACACTGT 207
Db 140 CAGTCTGTGACACGAGACTACAAATCACCCAGTCTCTCTCCACTGCTTACACTGT 199
Qy 208 CCGTCTTTTGTGACATTAACAAATGGCCGAGATGAGATTTCTTAAATCCG 267
Db 200 CCGTCTTTTGTGACATTAACAAATGGCCGAGATGAGATTTCTTAAATCCG 259
Qy 268 GAGTAATCAAACTTATTTATTTCTTAAAGACACAGTCA 308
Db 260 GAGTAATCAAACTTATTTATTTCTTAAAGACACAGTSA 300

RESULT 8
US-09-303-524A-1
; Sequence 1, Application US/09303524A
; Patent No. 6238873
; GENERAL INFORMATION:
; APPLICANT: CHAMBERS, JONATHAN K.
; APPLICANT: STEWART, BRIAN R.
; APPLICANT: AMES, ROBERT S.
; APPLICANT: SARAU, HENRY M.
; APPLICANT: FOLEY, JIM
; APPLICANT: ARNOLD, ANNE ROMANIC
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
; TITLE OF INVENTION: ANTAGONISTS OF THE INTERACTION BETWEEN THE HUMAN KIAA0001
; TITLE OF INVENTION: RECEPTOR AND LIGANDS THEREOF
; FILE REFERENCE: GP50007
; CURRENT APPLICATION NUMBER: US/09/303,524A
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/083,957
; PRIOR FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```



```

Db 720 GATACCATTTGATGCTATCTGAGGTGTCAGCCATTTGGGACTCTTCGATGTAAG 779
Qy 517 TCTCTTGGGGGCTAGATTCCTCTGTGTGTCATCTGGGCAATGTTCTTACTCTTT 576
Db 780 CATACCTTACAGAGTTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
Qy 577 GCTTAACATGATCTGACCAACAGGACCCGAGACAGAGATGTGAAGAAAGCTTTT 636
Db 840 GCCAACATCATCTGACCAATGCTGACCCAGACAGAGAAATATCATATCAGCTCATA 899
Qy 637 CCTTAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 900 ACTTAAAGCTCTTGGGGGCTCAATGCTAGGCACTAGCTAGCTAGCTAGCTAGCT 959
Qy 697 TTTCTGATTAATTTCTTAATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 756
Db 960 GTTGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
Qy 757 CCGGTATACGTAAAGACGAGGGGTGAGTAAAGTCCCGAGAAAAGGTGAACGTCAA 816
Db 1020 CAATTCAGAGGCAATTCATAAGT--CAGTCAAGCCGAAAGCAAAACATACAGAG 1076
Qy 817 AGTTTCATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
Db 1077 CATCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1136
Qy 877 TCTTACACCTGAGCCAAACCGGAGATGCTTGTGACTGCACTGCTGAAATATCTGTT 936
Db 1137 GCCTTCACTTTAGTACATTAGACAGGCTTTAGATGATGATGATGATGATGATGAT 1196
Qy 937 CTATGGAAGAGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 996
Db 1197 TTACTGCAAGAAATACATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256
Qy 997 CATTTTTCCTTGTGCAACTCTTCAAGAAATTCCTGATA 1037
Db 1257 TTACTTTTATCTGTAGTCAATTTCAAGATGGCTGTCA 1297

```

```

RESULT 10
US-08-702-344-27
; Sequence 27, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-702-344-27

Query Match      10.9%; Score 129; DB 1; Length 1243;
Best Local Similarity 48.4%; Pred. No. 4,6e-28;
Matches 428; Conservative 0; Mismatches 445; Indels 12; Gaps 2;

Qy 142 GAACACAGCTGCTGACACAGACATCAAAATACCCAGGCTCTTCCACTGCTCA 201
Db 201 GACAAACAGTCTGCTCTTCTGCCCCAGTTTAAAGATCTGGAGCCATTCACGATTTT 260
Qy 202 CACTGCTGCTGTTTTTTGTTGACTATACAAATGGCCGAGTGAAGATTTCTTTCA 261
Db 261 TTAATTTAGTTTCTGTTGGAATTTATGGAATTTGTTGCAACCTGGCTTTTATCA 320
Qy 262 AATCCGAGTAATCAACT--TTATTTATTTCTTAAGACACAGCATTTCTGATCT 318
Db 321 GAAGAAATACGAATCACAGGTGTGTGACATCTATTAAATTAATTTGCTTACACCGATT 380
Qy 319 TCTCATGATTCGACTTTTCCATTCAAATTTCTAGTATGCCAAATCGGAGACAGACC 378
Db 381 CTGCTTACTCTGCACTTACCAATGAAATTTGTTGACTGGGTGGCCACTTTGGAA 440
Qy 379 ACTGAAACTTTTGTGTGTCAGTACCTCCGTCATATTTTATTTTCAATATATATCAG 438
Db 441 GCTGAAGATATTCACAGTACCAAGTACAGCCCTCCTCATCTATATCAATATGATTTATC 500
Qy 439 TATTTATTCCTGGGACTGATTAATCTGATGCTTACAGACACACACAGCCATTAA 498
Db 501 AATTATCTTTTGAATTTGTACAGATTCACCGCTGCTTACACTGACACACAGCTGCA 560
Qy 499 AACATCCAAACCCCAAAATCTCTTGGGGCTAAGTATCTCTGTTGATCTGGGCAAT 558
Db 561 GATCTACCGAATACAAAGAACCCGATTTGCCAAATATGATACACCTGTGTGCTAAT 620
Qy 559 CAGTCTTACTCTCTTGGCTAATGATTTCTGACCAACAGGACCCGAGACCAAGA 618
Db 621 GGTCTTCTTAATATGTTGCAATATGATTTCCATCAAAAGATCAAGGAAAGTC 680
Qy 619 TGTGAAGAAATGCTCTTCTTAAATCAGAGTGGCTAGTGTGCAATGAATATGATA 678
Db 681 AATATGAGGTGTATGAGTTTAAAGGAAATTTGGAAGAAATTTGCAATTTGCTGACAA 740
Qy 679 TTACATCTGCAAGTATTTCTGATTAATTTCTTAATTTGATTTATGTTATATACAT 738
Db 741 TTTCATATGTGTAGCAATATTTTAAATTTCTGACCAATATTTTAAATTTCAATTTGCT 800
Qy 739 CATTAACAAAGAACTGACCGGTATACGTAGACAGAGGGGTAGTAAAGTCCCAAG 798
Db 801 TGTAAATTCGACAGCTCTACAG-----AAACAAAGATTAAGAAATTTACCAAAATG 851
Qy 799 GAAAAAGGTGAAGCTCAAAATTTTCATTTATCATTTGCTGATTTCTTATTTGTTTTC 858
Db 852 GAAAAAGGCTCTCATCAATCTTTAGTATGACACAGGGGTATCATATATGTTTCTTC 911
Qy 859 TTTCATTTTGGCCGAAATCTTACACCCGAGCAACCCGGGAGTGTGTTGAGCGAC 918
Db 912 TTACCAATTTGCGAATCCGATATACCTTCACGACAGACAGAAATATATCATGATGCTC 971
Qy 919 TGTGAAATATCTGTTATGTGAAGAGACACTGTGTTGTTACTTCTTAATATC 978
Db 972 AACCAAGATTTACTCTTCAAGCAAAAGAGGCTACAGTCTCTGCTGTGTGCAACT 1031
Qy 979 ATGCTGTGATCGTTCAATATTTTCTTGTGCAAGTCTTCAG 1023

```


TELEPHONE: 650-845-4170
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Liver
CLONE: 86700
US-08-911-320A-1

Query Match 6.5%; Score 77.2; DB 2; Length 1143;
Best Local Similarity 45.1%; Pred. No. 6.4e-13;
Matches 406; Conservative 0; Mismatches 488; Indels 6; Gaps 3;

QY 130 CTCGCGCCCTGGGACACAGTCTGTGACAGACTACAAATCCAGGTCCTT 189
DB 111 CTCAGCACACCAAAATGTTACTACCTGCTCCAGATGAAATGCTATCTGTT 170
QY 190 CCCAGTCTCTACACGCTCTGTTTGTGACCTTATCAAAATGGCCGATGAG 249
DB 171 AACCAATCTCTACTGTATTTTTCATCGTGGAGCTGTTGGGAACATATCGCCCTTA 230
QY 250 GATTTTCTTCAATCCGAGTAATCAACTTTAT--ATTTTCTTAAAGACACAGT 306
DB 231 TGTATTTCTGGGTATTCACCGTAAGAAATTCATCAATTTATCTACTTTAACGTAGC 290
QY 307 CATTTCTGATCTCTCATGATGATGATTTTCATTCATAAATCTTGTATGATGCAACT 366
DB 291 CATTTGAGACCTCTCTACTATCTTCTGCTCCCTTCGATTAATGATATATTAACCA 350
QY 367 GGGACAGACACAGTGAAGCTTTTGTGTCAGATTACCTCCGTATTTATTTTAC 426
DB 351 AAACAAGTGGACACTAGTGTGTATTTGTGCAAGTTGTGGAACTGGTTTATATGAA 410
QY 427 AATGTATATGATATTTCTCTCTGGAGTATATGATGATGATGATGATGATGATGAT 486
DB 411 CATGTACATTTGATGATTTTGTGATTTTGTGATTTTGTGATTTTGTGATTTTAA 470
QY 487 CAGGCAATTTAAACATCCAAACCCAAATCTTTGGGGGTATGATTTCTCTGTTGT 546
DB 471 TCGGTCTATACACAGGAGGCAATTAACACCAAAAGTA--TTTATGCTGTGTGT 528
QY 547 CACTGCGGCAATGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
DB 529 ATAGTATGATGCTGCTGCTGTGTGT--GGATTCTAATATATATTTTAACTTAAGA 587
QY 607 GAGAGACAAATGTGAAGAAATGCTTTCTTAAATCAGAGTTCGCTCTAGTGTGCA 666
DB 588 AGAGGAGCAATTAATTCACCAATGTTTTCATTCAGAGATTAACCAAAAGGAGA 647
QY 667 TGAATAGTAATTAATCATCTGTCAAGTATTTCTGTGATTAATTTCTTAATTTGTTAT 726
DB 648 AGCAATTTTAACTTATCTTGTGTGATTTCTGTGCTATTTTCTTAAATTAATCTT 707
QY 727 AATGTATACATCTTATCAAAAGAACTGTACCGGTATACGTAAGAAAGAGGGGTAGG 786
DB 708 TTTCAATATTAAGATTGGGAAGAAATCTAATGAGATTTCTTAAAGAGGTCAAAATTTCC 767
QY 787 TAAAGTCCCAAGAAAAGGTGAAGTCAAAAGTTTCTTATATCATGTCTATTTAT 846
DB 768 TAAATCTGTAAATTAATGCACTACCTGTAACCTCTTATTTGTTACTTATCAATTTTAC 827
QY 847 TTTTGTGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 906
DB 828 TATATTTGTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 887
QY 907 CTTTGAATGCACTGTGAAGAAATGCTGTTCTATGTAAGAGAGAGAGAGAGAGAGAG 966
DB 888 ATCATTTTGTCTACTGTGAAGAAATTTGTCAAAACCAATGATGATGATGATGATGAT 947

QY 967 TTCTTAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1026
DB 948 ATCTTAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1007

RESULT 13
US-09-217-101-1
Sequence 1, Application US/09217101
Patent No. 6143870
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
TITLE OF INVENTION: THROMBIN RECEPTOR HOMOLOG
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
City: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/217,101
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/911,320
FILING DATE: August 14, 1997
APPLICATION NUMBER: 08/467,125
FILING DATE: 6-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0041-1 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-845-4166
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Liver
CLONE: 86700
US-09-217-101-1

Query Match 6.5%; Score 77.2; DB 3; Length 1143;
Best Local Similarity 45.1%; Pred. No. 6.4e-13;
Matches 406; Conservative 0; Mismatches 488; Indels 6; Gaps 3;

QY 130 CTCGCGCCCTGGGACACAGTCTGTGACAGACTACAAATCCAGGTCCTT 189
DB 111 CTCAGCACACCAAAATGTTACTACCTGCTCCAGATGAAATGCTATCTGTT 170
QY 190 CCCAGTCTCTACACGCTCTGTTTGTGACCTTATCAAAATGGCCGATGAG 249
DB 171 AACCAATCTCTACTGTATTTTTCATCGTGGAGCTGTTGGGAACATATTCGCCCTTA 230
QY 250 GATTTTCTTCAATCCGAGTAATCAACTTTAT--ATTTTCTTAAAGACACAGT 306
DB 231 TGTATTTCTGGGTATTCACCGTAAGAAATTCATCAATTTATCTACTTTAACGTAGC 290
QY 307 CATTTCTGATCTCTCATGATGATGATTTTCATTCATAAATCTTGTATGATGCAACT 366
DB 291 CATTTGAGACCTCTCTACTATCTTCTGCTCCCTTCGATTAATGATATATTAACCA 350

```

Oy 367 GGGAGACAGGACCACTGAGCACTTTGAGTGCAGAGTTACCTCCGTCATATTTTATTTAC 426
Db 351 AAGACAGTGCACCTAGGTGTATTTGTGCAAGGTTGGAGAACACTGCTTATATATAA 410
Oy 427 AATGTATATCAGTATTTTCAATTCCTGTGGAGCTGATTAATCATCGATCCACAGAACACC 486
Db 411 CATGTACATTTAGCATATATTTTCCCTGGATTCATCATGTTGGATCCGTATATATAATTA 470
Oy 487 CAGGCCATTTAAAGCATCCAAACCCCAAATATCTGTGGGGGCTAAGATTCCTGTGTGT 546
Db 471 TCGGCTATACAGCAAGGAGGCAATTAACACCAAAAGTA--TTATGTCTGTGT 528
Oy 547 CATTGGGCATTCATGTCTTACTCTCTTTGCTTGGCTTAACATGATTTGACCAAGCAGCACCC 606
Db 529 ATAGTATGATGATGCTGTGCTGTGTG--GGATTCCTAATGATGATTTTAAACCTTAATAA 587
Oy 607 GAGAGACAGAAATGTGAGAAATGCTCTTTCTTAAATCAGAGTTGGGTCTAGTGTGCA 666
Db 588 AGGAGGCGCATATTTCCCAATGTGTTTCCATTAACAGATAAGCAATTAACGCAAAAGAGGA 647
Oy 667 TGAATATGTAAATTAATCATCTGTCAGTCAATTCCTGGATTTTCTTAATGTATATGT 726
Db 648 AGCCATTTTAACTTCACTTCTTGTGTGTAATGTTCTGCTAATTTCTTACTATATATCT 707
Oy 727 ATGTTATACACTCATTTACAAAGAACTGTACGGTCAATACGTAAAGACGAGGGGTAGG 786
Db 708 TTCAATATTAAATGTGGAGAAATCATATGAGAGATTTCTAAAGAGAGTCAAAATTTCC 767
Oy 787 TAAAGTCCCAAGAAAAGAGTGAAGCTCAAGTTTCATATCATATGCTGATTTCTTAT 846
Db 768 TAAATTCGTGAATATATGCCACTACACTGCTGAATCTCTTATATGACTTATCATTTTAA 827
Oy 847 TTGTTTGTCTCTTCCATTTTGTCCCGGAATTCCTTACACCTCGAGCAAAACCCGGAGT 906
Db 828 TATATGTGTGGTCCCATCATCATGCTTGTGATTCATCTACATTTCTCAACACTAAATGT 887
Oy 907 CTTTGACTGCACATGCTGAAAATATCTGTTCTATGTGAAAGAGAGCACTGTGTGTTAAC 966
Db 888 ATCATCTGCTACTGGAAGAAATTTTTCACAAACCAATGAGATCAATGCTGTCTCTC 947
Oy 967 TTCCTTAAAGCAAGCTCGATCCGTTCAATCTATTTTCTTTGCAAGTCCCTCAGAAA 1028
Db 948 ATCTTCAATAGTGTGTTAATCATCATGATCTATTTCCGTGATGTCCAGTACATTCGCAA 1007

RESULT 14
US-08-513-974B-353
: Sequence 353, Application US/08513974B
: Patent No. 6114139
:
: GENERAL INFORMATION:
: APPLICANT: Hinuma, Shuji
: APPLICANT: Hosoya, Masaki
: APPLICANT: Fujii, Ryo
: APPLICANT: Ohtaki, Tetsuya
: APPLICANT: Fukusumi, Shoji
: APPLICANT: Ohki, Kazuhiko
:
: TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
: TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
: NUMBER OF SEQUENCES: 380
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PC-DOS/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/513,974B

```

```

1 FILING DATE: 14-SEP-1995
2 CLASSIFICATION: 536
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: PCT/JP95/01599
5 FILING DATE: 10-AUG-1995
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: JP 7-093989
8 FILING DATE: 19-AUG-1995
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: JP 7-057186
11 FILING DATE: 16-MAR-1995
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: JP 7-007177
14 FILING DATE: 20-JAN-1995
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: JP 6-326611
17 FILING DATE: 28-DEC-1994
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: JP 6-270017
20 FILING DATE: 02-NOV-1994
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: JP 6-236357
23 FILING DATE: 30-SEP-1994
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: JP 6-236356
26 FILING DATE: 30-SEP-1994
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: JP 6-189274
29 FILING DATE: 11-AUG-1994
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: JP 6-189273
32 FILING DATE: 11-AUG-1945
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: JP 6-189272
35 FILING DATE: 11-AUG-1994
36 ATTORNEY/AGENT INFORMATION:
37 NAME: Resnick, David S.
38 REGISTRATION NUMBER: 34,235
39 REFERENCE/DOCKET NUMBER: 45753
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 617-523-3400
42 TELEFAX: 617-523-6440
43 INFORMATION FOR SEQ ID NO: 353:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 414 base pairs
46 TYPE: nucleic acid
47 STRANDEDNESS: double
48 TOPOLOGY: linear
49 MOLECULE TYPE: CDNA
50 FEATURE:
51 NAME/KEY: CDS
52 LOCATION: 28..384
53 US-08-513-974B-353
54
55 Query Match 6.4%; Score 75.6; DB 3; Length 414;
56 Best Local Similarity 50.0%; Pred. No. 1,1e-12;
57 Matches 189; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
58
59 Db 467 GATCGCTACCAAGAACACACACAGGCAATTTAAACATCCAAACCCCAAAATCTCTTGGG 526
60 19 GACCGGTATTTCAACATTTGTGAAGCCCTTTCACAGCTCTTCATCCACAGCTGTGAACATAC 78
61
62 Oy 527 GCTAAGATTCTCTGTTGTGTCATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATG 586
63 79 AGCAAACTGCTCTGCTGTGTGTTGTTGTCATGCTCTCTCCGCGCTCCCAACGTC 138
64
65 Db 587 ATTTGACCAACAGCAGCCGAGCAACAAGATGTGAAGAATGCTTTCCTTAATCA 646
66 139 ATTCTACCAACCAAGAGATTAGACACGTCGACGCAAGATTAATTCATGAGACTTAAAC 198
67
68 Oy 647 GAGTTGGTCTAGTGTGCATGAATAGTAAATTACATCTGTCAAGTCATTTTCTGATT 706
69
70 Db 199 GAATCGGGCCGCGACACAGCGTCACAACTCATCTTGTGGGCAATTTTCTGGCT 258

```

```

QY 707 AATTCTTAATGTGTATGTATGTATACATCTATTACAAAAGACGTACCGTAC 766
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 GGTCTCTTTTGTGTAATCATCTTTCTACACGCTATCACCGAAGAAATCTTAAGTCCAC 318
QY 767 GTAGAAGCAGGCGGTAGTAAAGTCCCGAGGAAAAAGGTGAACGTAAGTTTCAT 826
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 319 CTGAATTCGAAAGAAATTCATCTGGGTCAAAAAGAAATCTAGCCGCAATCTTCAGC 378
QY 827 ATCATGTGTATCTTT 844
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 379 ATCGTGTATCTCTCT 396

RESULT 15
US-08-513-974B-51
; Sequence 51, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohlaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273

```

```

; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-513-974B-51

```

```

Query Match 6.0%; Score 71.4; DB 3; Length 357;
Best local similarity 50.1%; Pred. No. 1.7e-11;
Matches 177; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 479 AAGACCAAGCCCATTTAAACATCCACCCCAAAATCTTGGGGCTAAGATTCTC 538
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 AAGATTGTGAAGCCCTTCTCAGTCTCTTCATCCAGTGTGAACATACAGCAATCTGTC 63
QY 539 TCTGTGTATCTGTGGCATCTGTCTTCTCTTGGCTTAACATGATTGACCAAC 598
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 TCGCTGTGTGTGTGTTGTCATGCTCTCTCTGCGGCCCAACATCTTCTGACCAAC 123
QY 599 AGGCAGCCGAGAGACAAAGTGAAGAAATGCTCTTCTTAAATCAGATTGGTCTA 658
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 CAGAGATTGAAGACGACGACAGATAAATGCATGGAATTAAGCAAGTGGCCGC 183
QY 659 GTCGTGATGAATTAAGTAATTAATCTGCAAGTATTTTGGATTAATTTCTTAAT 718
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 CAGTGGCAAGGCGTCAAACTACATCTTTGTGGGATTTTGTGCTGTGTTCTTTG 243
QY 719 GTTATTGTATGTATACATCTATTAACAAAGAACTGACCGGTACATAGTAAGAGAG 778
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 244 CTATATCTTTTCTACACGTGATACACAGGAAATCTTTAAGTCCCAAGCAATTCACA 303
QY 779 GGTGTAGTAAGTCCCAAGAAAAAGTGAAGTCAAAATTTTCAATTATCAT 831
    || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 304 AAGATTCCATCTCGGTCAAAAAGAAATCTAGCCGCAACATCTTCAGCATCGT 356

```

Search completed: December 11, 2002, 07:41:01
 Job time : 82 secs